Psychological Bulletin

CONTENTS

Psychological Bulletin

THE PRESENT STATUS OF THE KÖHLER-WALLACH THEORY OF SATIATION

HERMAN H. SPITZ[1]

Edward R. Johnstone Training and Research Center

In his book, *Dynamics in Psychology* (48), Köhler introduced an electrochemical wave theory to account for certain perceptual processes. He pointed out that the phenomena of stroboscopic movement and of the grouping of similar objects indicate that a percept process must be physiologically represented beyond its own locus. Using the physicist Faraday's investigations as a frame of reference, the primary postulate set forth by Köhler was that a percept is a dynamic agent which extends into the surrounding brain tissue. Taking reversible figures as an example of a percept process in a more active state, and pointing out that these reversals increase in speed with increased number of observations, Köhler concluded that: " . . . a figure process seems to have some effect by which it tends more and more to block its own way," (p. 72) much as in electrolytical conduction local resistances may be heightened and the permeability of the interfaces themselves altered. Since the spontaneous transformation of certain percept patterns is a behavior analogous to that of currents in electrolytes, the assumption was made that percepts are actually associated with electric currents in the nervous system. It

was further postulated that a wave action occurs through the medium surrounding the neurons.[2]

In order to test in what way the prolonged occurrence of a percept process in a given area changes the medium in which it is located, Köhler turned to figural aftereffects as demonstration phenomena. Figural aftereffects may be elicited in the visual, kinesthetic, and even auditory sphere; but early investigations were primarily in the visual area. The figures used may be any shape, but to demonstrate a typical experimental procedure the following example from Köhler's early studies (48, p. 88) will be used. The initial stimulus field consists of an Inspection Figure, in this case a single black outline circle on a plain white ground. A measured distance below this circle is a small dot or cross, called the Fixation Point, on which the subject is told to fixate. Throughout this inspection period the circle is in peripheral view. After a measured period of time this stimulus field is replaced by the Test Figure which has its own Fixation Point with objectively equal sized squares symmetrically above and below it. The Test Figure is constructed so that when its Fixation Point is fixated the

---

[2] Köhler describes the physiological process in some detail in his book (48) and in a later article with Wallach (56).

upper square falls *within* the area previously occupied by the circle of the Inspection Figure. The subject is immediately asked to compare the two squares and will usually report that the upper square is smaller than the lower square. The aftereffect of the inspection circle is revealed in the apparently decreased size of the upper square. The bottom square of the Test Figure, unaffected by any previous figure, is present only as a comparison square.

In a follow-up article with Wallach (56), the word "satiation" was used as a descriptive term for the alteration in the brain medium brought about by the prolonged presence of a given figure in a particular cortical area. In other words, the figural aftereffect is the manifest measurement of a physiological process labeled "satiation." Continued presence of figural currents increases the resistance of the tissues through which they flow, and an area of resistance or impedance is created in the neighborhood of the contours of any fixated figure. The contours of the test figure will then recede from the area of impedance created by the contours of the inspection figure.

According to satiation theory, then, the explanation for the apparent decrease in size of the upper square in our example lies in the fact that the contours of the square were encompassed by the contours of the previously fixated circle and therefore receded from the area of impedance, making the square appear compressed. Test figures may exhibit changes in size, luminosity, depth, and displacement—effects which appear to be independent of each other (51).

As noted above, many types of figures may be used to demonstrate the aftereffect, although solid figures are usually avoided because of the disturbing afterimage. The effect can be elicited with switching monocular viewing, indicating its central origin. Besides presenting the results of research with these figures as corroborative evidence of their hypotheses, Köhler and Wallach termed the Gibson curved line effect (31) a special case of figural aftereffects to be subsumed under the same general physiological law.

Since the original publications by Köhler and Wallach there has been a good deal of attention paid to the figural aftereffect and to the theory it is purported to support. Many of these studies are corroborative, but some contradict each other and the original articles, and some lead to the presentation of new theories or alternate explanatory hypotheses. It is the purpose of the present review to pull together all the research on satiation theory. Special attention will be paid to the apperception of the brain-damaged, since it is so often the case that we reach an understanding of normal processes by an investigation of deviation from normality, and because it is the present reviewer's belief that satiation theory may have much to contribute to an explanatory theory of organic apperception. Finally, it is hoped that an understanding of the present status of satiation theory will open new avenues of approach in the area of perception and perhaps contribute to what is the ultimate outcome of such research: an adequate brain model and the reduction to fewer and fewer general laws of the seemingly diverse types of phenomena given to us through our sense modalities.

METHODOLOGY

This section will review the problems encountered and the methods

employed in attempts to demonstrate the presence or absence of satiation. The most frequent procedure has been to use the visual figural aftereffects; but kinesthetic figural aftereffects and ambiguous and reversible figures, as well as apparent movement phenomena, have also been frequently employed.

General Corroborative Studies

Shortly after the appearance of the Köhler and Wallach article, Walthall (91) used college students to study various aspects of what he termed the "Köhler effect." Among other things, he reported that smaller figures give the best effect and that measurements can be made to within one-sixteenth of an inch. He concluded that the Köhler effect is real and stable enough for measurement.

In his attempt to verify the statistical frequency of the figural aftereffect, George (29), using only one inspection figure to correspond with one of his two test figures, found his results to be inconclusive. In fact, four of his subjects reported a change in the direction opposite to what would be expected from Köhler and Wallach's results. However, when he used two inspection figures, one for each test figure and each working in the opposite direction from the other, his results were completely in accord with Köhler and Wallach's findings.

The Effects of Color and Brightness

Köhler (49, p. 240) stated that, although he obtained figural aftereffects when the object and its environment had the same brightness and differed only as to color, he was under the impression that the effects tended to be weaker than those obtained when the figures used were strikingly different in brightness. In a later study (55), however, he indicated that a medium brightness difference seemed to be optimal.

Hochberg and Triebal (42) tested the effect of colored stimuli on figural aftereffects and found that chromatic differences can neither modify nor generate the phenomenal aftereffect. However, when the figures differed from the ground largely or only in chroma and not also in brightness, no measurable figural aftereffect was found, although the forms of the figure were perfectly clear. These results clearly suggest that brightness differences between figure and ground are necessary for the physiological processes presumably responsible for figural aftereffects.

Walthall (91) reported that a white ground produced the greatest amount of fatigue and discomfort, while a manila ground was more comfortable and did not lessen the effect in any way. Marquart (67), addressing herself to a somewhat different problem, noted that when more light fell on one figure then on another (as for instance from a window on one side of the room), it had little effect on the aftereffects produced.

Freeburne and Hamilton (27), following Walthall, tested the effect of the brightness of the stimulus field on figural aftereffects. They found that differences in brightness did not influence the average number of aftereffects, although the number of aftereffects produced on a high brightness field was initially higher and then tended to decrease with successive presentations. The opposite occurred when the field was of low brightness. These results suggest that the effect will be more immediate if the stimulus is presented on a bright field and that "over-satiation" may cause a reduction in the aftereffect due to the spread of satiation,

a phenomenon later expanded upon by Köhler and Fishback (**52, 53**).

Not only does the brightness of the figure and ground influence the aftereffect, but one of the phenomena consequent upon previous fixation is said to be a change in brightness level, a fading away or graying of the test figure. Bevan (**7**) devised a unique method of exploring this effect by projecting identical pairs of dots or small outline figures on a screen in such a way that one fell within the area previously occupied by the inspection figure, and the other, outside. These pairs of test figures were presented subliminally and increased in luminosity until the threshold of perception was reached. Out of 233 judgments in which one of the pair of test figures was seen first, in 67% of the judgments the figure falling outside the satiated area was seen at a lower illumination than was the inside figure. The outside figure was reported first in a majority of trials by 80% of the subjects. Furthermore, the test figure with the lower threshold was frequently reported as being darker, larger, or nearer than its partner, in different combinations by different subjects, corroborating Köhler's descriptions and the apparent independence of the various aftereffects.

Temporal Factors

Hammer (**35**) reported that the displacements of the figural aftereffects are at a maximum immediately after the inspection period and decrease gradually to approach zero at about 90 seconds. Considering the length of inspection time, she found a recognizable effect after as little as 5 seconds, with a maximum reached at 60 seconds. She noted that the results were similar to those of Gibson and Radner (**33**) on the tilted line effect

as a function of inspection time, despite the fact that the stimulus situations were different.

Using three-minute fixation periods and solid black rectangular figures, Krauskopf (**58**) also found that the size of the aftereffect decreased as the exposure time of the test figure increased. As Krauskopf points out, these findings contradict Köhler's statement that prolonged inspection actually reduces the impedance of the corresponding cortical area and therefore allows the subsequent test figure, operating in a more permeable field, both to create its own impedance and to be affected by it (**49**, p. 233). This statement by Köhler, which Krauskopf uses as a frame of reference, is rather puzzling, since it is contrary to the usual description of satiation as an increased impedance created by the prolonged presence of a stimulus figure, with subsequent figures moving away from the area of impedance. Using this latter definition, the results found by both Hammer and Krauskopf are easily explainable, since the amount of impedance should gradually dissipate.

When both the test and the inspection figures are presented simultaneously the results are exactly opposite those obtained when the test figure followed the inspection figure, thus indicating that temporal delay is necessary for obtaining the figural aftereffect (**29**).

Directional Factors

Köhler and Wallach (**56**) found that the figural process is more intense between vertical than between horizontal lines and is generally stronger and appears earlier in the lower half of the visual field. However, Weitz and Post (**95**) in their stereoscopic studies could not cor-

roborate Köhler and Wallach's re-
sults. After satiation, none of their
10 female subjects reported the test
figure in the lower half as smaller, al-
though 9 out of 10 had previously re-
ported the expected aftereffect when
the upper figure was used. Even
after no inspection figure was used,
the lower square appeared larger to
half the subjects, and brighter to 8.
Removing the lenticular prisms did
not change the results.

Marks (65), however, reported re-
sults contradictory to those of Weitz
and Post. With appropriate inspec-
tion figures, the lower rectangle ap-
peared smaller to 18 out of 20 sub-
jects, and larger in only one case.
Furthermore, the depth and lumi-
nosity changes after satiation oc-
curred only in the lower portion,
while the size effect occurred to a sig-
nificant degree in all fields. He con-
cludes that the Köhler effect is per-
haps present only in the judgment
of the size variable, and even here it
is not universal.

Hammer's (35) subjects used a 14-
inch monocular viewing tube, 1 inch
in diameter, to view test and inspec-
tion figures which were both on the
left side of the fixation point. She
found a larger displacement of those
figures which were on the far left of
the fixation point. Her explanation
of these results is that the right-hand
figure was closer to, and interacted
with, the fixation point, thus reduc-
ing the effectiveness of that figure in
inducing a figural aftereffect. But
Gibson (31) had previously noted
that the aftereffect is slightly larger
when the inspection figure is exposed
peripherally rather than foveally, and
one wonders if there might not be
some increase in the measured after-
effect with increase in distance from
foveal fixation. On this point, Bender
and Teuber (6) reported that defects

in the critical flicker frequency of
brain-injured subjects were more evi-
dent when tested in the periphery of
vision.

Using a reversible perspective fig-
ure, Carlson (14) found vertical fluc-
tuations more clear-cut than hori-
zontal ones. George (29) found no
significant difference in establishing
an aftereffect with horizontal or verti-
cal placement of figures, but there
was no gradient of response choice in
his experiment. Subjects merely re-
ported whether or not they perceived
the aftereffect. Although George,
whose subjects used alternate monoc-
ular viewing, reported that there was
no significant relationship between
eye dominance and the size of the
aftereffect, Marquart (67), whose
subjects viewed the stimuli binocu-
larly, found that the aftereffects were
significantly greater on the side of
the non-dominant than on the side of
the dominant eye.

Frequent differences have been
found in the apparent size of test fig-
ures in opposite halves of the visual
field when no inspection figures were
used. Brown (12) labeled these dif-
ferences "half-meridional differences"
(HMD's) and disputed the statement
by Köhler and Wallach that inho-
mogeneities in the visual field play
only a minor role. In fact, HMD's
enter into changes in the Müller-
Lyer illusion and therefore must al-
ways be measured and taken into ac-
count when measuring figural after-
effects. The difference between pre-
and post-inspection measurements of
the test figure will result in the truest
measure of the satiation effect.

George (29) reported that only 4
out of 20 subjects saw the test squares
as equal when they were presented
initially and without previous stimu-
lation, but judging from Marks' (65)
results this pre-inspection or control

measurement might actually enhance the production of the aftereffect. Furthermore, Marks found that more of his subjects perceived the objectively equal squares as being equal when the control readings were given after instead of before the experimental situation. He attributed these results to a sharpening of the subjects' discriminatory ability.

Distance Between Figures

Köhler and Wallach (56) reported that a distance of about one-quarter inch between inspection and test figures was optimal, and that within certain limits the size effect decreases as the outlines of the inspection figure approach or recede from those of the test figure beyond this optimal distance. They called this phenomenon the "distance paradox." When the two figures coincide there is no spatial displacement, although the test figure may appear paler and farther back, and even diminished in size.

With solid black rectangles as inspection figures and dots as test figures, Fox (25) verified the existence of the distance paradox and found that the maximum aftereffect occurred when the boundaries of inspection and test figures were separated by between one-quarter and three-quarters of an inch. However, in contrast to Köhler and Wallach's findings, a test dot which coincided with the border of the inspection figure was in fact displaced. Another contradictory finding was the displacement of the test dot *toward* the nearest line of the inspection rectangle, when it was encased within that rectangle, a finding analogous to the attraction phenomenon reported by Smith (82) and the adaptation phenomenon reported by Gibson (31) and by Prentice and Beardslee (76). Unlike Smith, however, Fox does not suggest throwing out satiation theory

completely. Rather, he postulates that adaptation of the asymmetrical inspection figure to the norm combined with satiation to produce his results.

In order to test this postulate, Fox eliminated the effects due to adaptation by using symmetrical stimuli. Under these conditions he found that not only were displacement aftereffects smaller than when asymmetrical inspection figures had been used, but also that the two contradictory findings of his first experiment were eliminated. He concluded that in his experiment the effects of adaptation to a norm seem to be as powerful as those of satiation, although the interaction between the two is not necessarily of an algebraically additive nature.

Walthall (92) performed an interesting experiment in which he varied the distance of comparison test circles from a satiated area. He reasoned that the current flow stemming from *outside* the inspection circle would compress the comparison circles, although to a lesser degree than the enclosed test circle, since the side of the comparison circle nearest the current flow would be displaced more than the other side of that same circle. Because the effect dissipates with distance, those comparison circles at a greater distance should appear larger than circles closer to the satiated area. His results, also achieved with switching monocular viewing, were consistent with his hypothesis, indicating that field strength varies as a function of distance within the field.

Using his luminosity thresholds as measuring criteria, Bevan (7) found that the strength of the figural aftereffect varied with the distance between inspection and test figures. The aftereffect was greatest when the loci of the test and inspection figures

coincided, and diminished as the two loci increased in distance from each other. The differences at five successively increasing distances were all reliable, and the results were always most pronounced in the lower half of the visual field. These results are not inconsistent with the distance paradox since the graying effect is said to be independent of the displacement effect.

Distance Between Observer and Figures

Prentice (74) initially reported that the figural aftereffect maintained its size independent of the distance from the observer, but later (75), following in design an unpublished experiment by Thouless, contradicted his initial findings. His second arrangement was such that the inspection figure was retinally smaller than the appropriate test figure but because of distance and size constancy appeared larger to the subject. Of 65 subjects, binocularly fixating for 40 seconds, 40 reported that the appropriate (satiated) test figure appeared larger than an objectively equal but nonsatiated comparison figure. This result is compatible with the theory that the visual angle rather than the apparent size is fundamental to the dynamics of figural aftereffects. Hochberg and Bitterman (41) also found evidence in favor of this latter conclusion. Sutherland (85), however, proposed that both apparent size and retinal size are operating in producing figural aftereffects. He presented confirmatory results from experiments in which he alternately eliminated as much as possible the influence of retinal size and of apparent size.

Changes During Fixation or After Scanning

Weitz and his associates (94, 95), in the course of their stereoscopic studies, noted that their subjects reported some peculiar effects during the initial inspection period. Some changes occurred after as little as two minutes of fixation. Square figures seemed to be more stable as inspection figures, but the same type of change occurred in the circle without previous fixation as occurred in the square either with or without previous fixation. Marks (65) reported a significant incidence of the Köhler effect in situations where the appropriate inspection figures were not present.

Following up the observations of Weitz and Post and his own previous findings, Marks (66) studied the phenomena reported by his subjects during a 4-minute fixation period. He categorized the reports of 60 college students, who observed monocularly, into the following four areas: movements or changes in the shape of the fixation point, a periodic variation in light intensity or differential spatial intensities at any given moment, partial or complete blotting out of the phenomenal field, and a tendency for the parafoveal circle to assume various elliptical or polyangular shapes. Although he raises the possibility that certain of the phenomena elicited—if they are central in origin—may be likened to ambiguous figures and explained on the basis of maximal polarization and consequent displacement, he nevertheless concludes that it may be necessary to modify field theory to include the influences of oculomotor fatigue, retinal adaptation, and other such peripheral factors. In any case, the results are inconsistent with Gestalt principles of closure and good form.

The work of Marquart (67) also indicates some inconsistency between the Gestalt principle of prägnanz and the theory of satiation. She reasoned that, if perception of a prägnant fig-

ure involves less kinetic energy, as Gestalt theorists propose, then this type of figure should produce a smaller figural aftereffect than a complex figure. However, the opposite effect occurred, leading the author to the conclusion that a new theory of figural aftereffects is needed.

Luchins and Luchins (62, 63) attempt to find some solution to the problem of how the fixation of a figure can overcome such established Gestalt principles as simplicity and closure. They suggest as a possible explanation that a tendency toward regularity in the condition of the cortical medium may in fact be phenomenally represented as irregularity. They add that, if this is the case, then any Gestalt principles of stability or simplicity would also require a stipulation of the cortical conditions under which stability or simplicity holds.

Prentice and Beardslee's (76) subjects were required to compare a periodically exposed test line, which was variously tilted, with a continuously exposed standard line which was at all times actually 10° off the horizontal in one exposure, or vertical in another. Their subjects tended, on the average, to equate the standard line with a test line that was 8° off the horizontal or vertical. This normalization of 2° was an effect independent of the previous presence of other contours in the appropriate part of the visual field. The effect is also apparent when the head is tilted, ruling out an explanation in terms of inhomogeneities of the primary visual projection areas. In their theoretical discussion of these results they suggest the extension of the concept of satiation to include developmental factors which give some representation to the quality of "verticalness" in the nature of the brain process.

In a recent paper, Held (38) reported that one minute of haphazard scanning of a patterned inspection figure was sufficient to establish aftereffects similar to those attendant upon steady fixation. The patterned inspection figures were a series of parallel lines and grill-like arrangements oriented in particular directions or with a particular curvature, quite different from the usual segregated square, circle, or oblong inspection figure. In one experiment, a patterned angle of 10° caused a greater tilt effect than smaller or larger angles, a phenomenon analogous to the distance paradox. The author concluded that it is the patterning of the scanned exposure field which is the critical factor, while the usual inspection figures merely restrict the range and locus of the aftereffect.

Kinesthetic Figural Aftereffects

Köhler and Dinnerstein (50) had their subjects rub the fingers of one hand along the sides of a narrow or wide strip of cardboard for one or two minutes and then use the same hand to judge a standard sized strip. After stimulation by the smaller strip, the standard strip was judged to be larger; the opposite effect occurred with initial stimulation by the larger strip. The authors felt that these results might ultimately be explained on the basis of satiation theory.

Nachmias (70) employed a kinesthetic apparatus more closely analogous to the stimulus figures used in producing visual figural aftereffects. His subjects placed one hand on a bar above or below a standard test bar. After 60 seconds, both hands were simultaneously placed on the standard horizontal test bar and the subject was required to estimate the level of the test hand as compared with the unstimulated hand. Four-fifths of his subjects reported the ex-

pected aftereffect. However, further experiments produced no evidence of a kinesthetic distance paradox analogous to that found in visual aftereffects.

After testing brain damaged patients and normal controls, Jaffe (43) reported that the absolute amount of the kinesthetic aftereffect for both groups was much greater after stimulation with a larger as compared with a smaller strip. Handedness had no influence on the aftereffect. Wertheimer (99), however, found that right handed subjects overestimated the width of a bar presented to their right hand and that right handed women did so to a significantly greater degree than did right handed men. Handedness seemed to be less pronounced in left handed subjects. This difference in handedness Wertheimer labeled a "bilateral kinesthetic difference" (BKD), and he agreed with Brown (12) that pre-inspection measurements must be made when measuring figural aftereffects. Without such control measurements the amount of kinesthetic figural aftereffect would have been about 27%, whereas with control measurements he found that the aftereffect distortion still measured a significant 13%. It is of interest to note that women, besides making higher control matches, produced a significantly smaller aftereffect then men, suggesting the possibility of sex differences in ability to satiate.

In another study, Wertheimer (101) found a statistically significant positive relationship between visual and kinesthetic aftereffects and also in changes of visual and kinesthetic figural aftereffects after changes in metabolic rate. This is corroborative evidence for the supposition that the figural aftereffects found in different sense modalities are produced by the same physiological process.

Auditory Aftereffects

Jones and Bunting (45) measured the auditory aftereffect by the method of limits. Their subjects determined the subjective midpoint before and after a two-minute exposure to a constant sound which the experimenter displaced to one side of the median plane. The subjective midpoint was found to be displaced in the direction of the sustained tone and the effect was greater when the continuous auditory simulation was set 60° from center than when placed 90° away.

Deutsch (19) hypothesized directly from satiation theory that a sustained tone in the auditory scale ought to separate the neighboring tones, increase the tonal interval between them, and thereby decrease the phenomenal difference limen (DL). Since in fact seven out of eight experimental subjects did show a significant decrease in DL, while a control group varied randomly, Deutsch concludes that there is something analogous to visual figural aftereffects in hearing. Because contralateral monotic stimulation was used with the experiment, the effect is said to be central.

A new deduction from satiation theory was made and auditorily tested by Krauskopf (57). The deduction was that satiation at the medial plane should reduce the spatial difference limen at this point. Using a horn on a horizontal track in front of his subjects, who were in total darkness, Krauskopf reported that the standard deviation of settings around the midpoint decreased following satiation. It decreased also in a nonstimulated control group, but to a significantly lesser degree. In a further experiment, the two minute tone was presented at varying distances to the side of the medial plane.

It was found that the displacements were toward the stimulated tone and increased as the distance from the medial plane increased. Despite the fact that no larger distances were tested, one of Krauskopf's conclusions is that the distance paradox was partially demonstrated. He bases this conclusion on the fact that a decrease in the size of the displacement had previously been shown for larger distances by Jones and Bunting.

Use of Reversible and Ambiguous Figures

Köhler and Wallach (**56**) reasoned that since prolonged presence of a figure in a given location tends to operate against further presence of this figure in the same place, then figure-ground relationships will suddenly be reversed when the critical level of satiation for one perception is reached. Although the direction of figural currents is reversed when the color of figure and ground are reversed, aftereffects should remain the same since satiation is nondirectional in nature. In other words, there is an equal amount of impedance no matter what the direction of the current.

Hochberg (**40**) questioned these latter assumptions. Using ambiguous black-white test figures, he found that going from a black satiation figure to a black test figure, or from white to white, brought only chance results whereas reversals to the opposite color yielded results compatible with satiation theory. This suggests that there may well be a weak directional effect in satiation.

Carlson (**14**) extended Hochberg's technique a step further by using a reversible perspective figure (a bar through a ring). He proposed that prior satiation of one alternative of a reversible perspective figure would depress that alternative in favor of another. That is, the nonsatiated alternative should appear first upon subsequent fixation of the reversible figure. Using male college students and only 15-second fixation periods, his results were consistent with his hypothesis and with the theory that satiation may underlie certain types of reversals. He followed with a second experiment, this time using a more meaningful ambiguous figure (wife and mother-in-law). Results here were again highly significant, but in the opposite direction, thus ruling out the possibility that suggestion was a factor in his initial results. Carlson accounts for the contrary results of his second experiment on the basis that the reversal here involved a change of meaning but not in apparent depth or perspective. He opens up an interesting area of speculation by suggesting that Köhler's satiation hypothesis, as it stands now, cannot be applied to the reversal behavior in meaningful figures. In other words, his results raise the question of how much and in what situations individual experience and personality impede or accelerate—either consciously or unconsciously—satiation effects. At any rate, "pure" satiation can probably best be tested with a stimulus that is as nonpersonal and "passive" as it can be made.

Brown (**13**) felt that if the same physiological process is responsible for both figural aftereffects and apparent change in ambiguous figures, then the rate of apparent change might be a more reliable means of measuring the rate of satiation. He used eight subjects; and his apparatus, viewed monocularly, produced a three-dimensional rotating effect. The rate of apparent change increased rapidly up to two minutes, then gradually leveled off. There were marked individual differences in all

recordings, but curves for the right and left eye were identical or very nearly identical for all subjects. He concluded that his findings were consistent with the hypothesis that the same physiological process is responsible for both figural aftereffects and apparent change in ambiguous figures.

Also using a moving, reversible figure, Fisichelli (24) reported that increasing the complexity and the speed of Lissajous figures enhanced the reversal rate; and he postulated that the enhancement was due to a concomitant increase in the rate of stimulation of a group of nerve cells in a given area. He concluded that results lend support to the assumptions of satiation theory without, however, constituting any direct proof.

Klein (46) reports evidence of the existence of a continuum from a "lag" or "leveling" attitude to a "sharpening" attitude. Taking the extremes of a group of 50 adult men and women, the lag group responded more slowly to change, had more difficulty finding imbedded Gottschaldt figures and hidden faces, and tended to diminish differences between figure and ground. Certain perceptual attitudes of "self-inwardness"—such as avoidance, passivity, and dependence —were associated with the lag group, as compared to the "self-outwardness" of the more energetic sharpening group. Klein feels that this analysis touches ground with research in time-error and, through it, with the study of satiation effects.

As one of three perceptual methods of testing the effects of artificially changed metabolism, Wertheimer and his associates (102) tabulated the number of reversals reported during a 30-second fixation of a Necker tube. Unfortunately the numerical results were not presented, although with decreased metabolism there was a significant drop in perceptual modifiability.

Apparent Movement

Crutchfield and Edwards (16) reasoned that, since individual subjects soon establish for themselves a prevailing amount and direction of autokinetic movement, then the prefixation of a crescent shaped figure on one side of the autokinetic light would set up an area of local cortical impedance which would affect the amount and direction of autokinetic movement. That is, if the prevailing direction of movement is to the right, prefixation of a crescent figure on the right—opposing the direction of movement—should slow it up or even change its direction. Prefixation on the left should increase movement. They found, however, a marked drop in the extent of autokinetic movement following fixation on *either* side and no significant variation in the prevailing direction of movement. When the fixation time was dropped from two minutes to one minute (20), there was a significantly greater decrease in movement when the direction of movement was opposed than when it was supported, but still the direction of movement was unaffected. The authors are at a loss for an explanation of their results but suggest that satiation must be only one of many variables affecting autokinetic movement.

Adding a variable to the work of Crutchfield and Edwards, Livson (61) designed his experiment to test the Gestalt theory of isomorphism. Also using the autokinetic phenomenon as a test figure, his inspection stimuli were four conditions of alternating flashing lights, only one of which created apparent (stroboscopic)

movement. He hypothesized that prolonged inspection of the apparent movement stimulus would reduce the amount of autokinesis to a greater extent than would similar fixation on stimuli which, because of changes in temporal sequence, produced such nonmovement stimuli as alteration and simultaneity. Further, the non-movement stimuli should not differ from each other in their effect on autokinesis. All the above hypotheses were confirmed. Livson concludes that these results show a relationship between the phenomenal field and its cortical correlate, lending support to the more general form of the isomorphic principle.

Deatherage and Bitterman (18) reported that the path of apparent movement across a previously fixated field was "bowed" as if the movement were circumventing the satiated area. When the inspection figure was placed laterally, completely separating the two points which stimulated the stroboscopic effect, succession was reported at a rate of alternation previously set for optimal movement. In some of their subjects an increased alternation rate reinstated the apparent movement, while in others there was simultaneity without intervening movement.

In essential agreement with Werner and Thuma, Shapiro (79) contended that the failure of the brain-damaged to report apparent motion was due to an exaggeration of inhibitory effects. He reasoned that he could induce inhibitory effects by the continuous stimulation of normal subjects who would then, like the brain-injured, show a deficiency in the capacity to perceive apparent movement. Using a tachistoscopic apparatus similar to that used by Werner and Thuma (97), with switching monocular viewing, he found a significantly raised

threshold of apparent motion after continuous stimulation. Although Shapiro explains his results primarily in terms of inhibition of irradiation and, following Bevan (7), notes that luminosity may have been lowered by previous stimulation, his results are certainly consistent with satiation theory. That is, continuous stimulation may set up an impedance (satiation) so that the perception of a satiated normal is like the perception of a patient with an injury (impedance) in the crucial area of the brain.

Brenner (8) reported that continuous stimulation decreased the range of apparent movement by causing simultaneity to be perceived at longer time intervals, and succession at shorter time intervals, even when her subjects were given different types of continuous activity—including visual and auditory stimulation, pacing up and down the room, and two minutes of simple arithmetic. Each type of continuous stimulation resulted in a highly significant and quite comparable decrease in range of apparent movement. These results contradict those of Livson and certainly argue against a point-to-point, isomorphic relationship. Brenner suggests that the neural process underlying the effects of continuous stimulation may possibly be localized in the preoccipital area, Brodman Area 19, a part of the visual system which can be affected by stimuli from all sense modalities.

Cross-Modal Study

Another study which challenges the isomorphic position of satiation theory is Jaffe's (44) cross-modal investigation in which large and small strips of white paper, visually presented, affected the kinesthetic judgment of a standard bar of aluminum. The kinesthetic standard appeared to

increase in width after visual fixation on a narrow strip of paper, and to decrease after fixation on a wide strip. The latter decrease achieved statistical significance. Because these same results were not obtained without use of a standard visual strip, the author feels that the important factor operating is some general frame of reference, possibly explainable in terms of Helson's theory of "adaptation level." If these results can be cross-validated, they would appear to be quite damaging to the theory of satiation, which posits a cortically represented spatial relationship between inspection and test figure.

Immediate Illusions

Summerfield and Miller (84) investigated the similarity between certain immediate visual illusions and figural aftereffects, and the possibility suggested by Köhler and Wallach that immediate self-satiation may be responsible for certain illusions. The figure chosen for the investigation was the "herring-bone" field which created the illusion that parallel lines were bent toward one another. In the aftereffect experiment, the field became the inspection figure and the parallel lines were used as the test figure. Results indicated that displacement was greater for the immediate illusion than for the aftereffect and that the illusion was significantly greater without fixation than with fixation. The authors interpret these results as indicating that some mechanism other than self-satiation is operating in the creation of illusions. However, the authors make no mention of the work of Köhler and Fishback (52, 53) on the Müller-Lyer illusion, and the fact that that illusion too is lessened by long fixation periods. In fact, we can assume that in the area of greatest

electrotonic activity two separate lines, when first perceived, will appear to be closer together, and only with fixation will satiation build up to a point where there is impedance.

Satiation and Metabolic Rate

Michael Wertheimer and his co-workers (100, 101, 102, 103) conducted a series of studies on individual differences in metabolic rate. Using figural aftereffects as a measure of metabolic efficiency, they worked on the premise that a relatively large figural aftereffect reflects high physiochemical modifiability and consequently a high metabolic efficiency. The word "efficiency" was used, since aftereffects were found to be maximal in the normal range of metabolic functioning and to fall off on either side.

Comparing their own work—based on satiation theory—with the Klein-Krech theory of cortical conductivity, the pivotal point of difference is the postulated curvilinear relationship between metabolic efficiency and perceptual modifiability. Whereas the Klein-Krech theory would predict an increase in figural aftereffects with a lowered metabolic rate (i.e., low basal cortical conductivity), Wertheimer would predict a decrease in the aftereffect, a prediction which was confirmed. Neither theory, however, could account for the lack of decrease in kinesthetic aftereffects and in reversible figures when metabolic processes were raised.

In a further extension, Wertheimer reasoned that a lowered efficiency of tissue metabolism in schizophrenics would imply less effective polarization. On the basis of results on a visual figural aftereffect test, this reasoning was experimentally supported. Interestingly enough, in view of his similar findings with kinesthetic

aftereffects, 12 female controls were less satiable than the male controls, although generally more satiable than the schizophrenics.

There would seem to be a need for correlative and theoretical studies comparing the above results with the results of Engel and Margolin (21) and of Davis (17). These latter authors studied the effects on the electroencephalogram (EEG) of differing carbohydrate metabolism and of changing blood-sugar levels. Similarly, Landis (59) found that drugs which decreased the availability of blood sugar and oxygen decreased the threshold of critical flicker frequency (CFF). An increased efficiency of vascular supply raised the CFF threshold.

EEG Studies

In order to investigate more directly the physiological correlate of pattern vision, Köhler and his associates (54, 55) made use of the EEG in very extensive studies measuring the effect on cortical currents of the perception of moving objects, stationary objects, after potentials, and audition. A full summary of their lengthy experimental findings and theoretical discussions is beyond the scope of this paper, but the authors found that the active parts of the visual areas seemed to be "surface positive" in relation to the less active parts. This would indicate that the stimulated area of the cortex is pervaded by currents which flow in corresponding directions, a result compatible with satiation theory. In the latter of the two articles, the authors indicate that the current within a circumscribed area will not only be more concentrated than it is in the surrounding tissue but the flow within and the flow without will be sharply segregated. That is, the flow within

is in the opposite direction to the flow without. The current flows through cortical objects and returns to these objects through their environment, and therefore the spatial continuity of cortical action is suddenly interrupted at the boundary line.

Animal Study

Lashley, Chow, and Semmes (60) have reported a comparative study of satiation theory. Those authors placed strips of gold foil, as well as gold pins, in contact with, or through, the visual cortex in the macular area of monkeys. Despite this, they found no noticeable disturbance of visual functioning. Since it was reasoned that metallic conductors should have altered the figure currents, the results of the experiment were presented as evidence contrary to satiation theory. They noted also that the theory is incompatible with known facts of neuroanatomy and pathology and makes no attempt to deal with the problem of stimulus equivalence or to explain how electrical fields influence the efferent neurons.

An Alternative Theory

Osgood and Heyer (72) are of the opinion that figural aftereffects may be accounted for in terms of more generally accepted neurophysiological principles. They draw heavily upon Marshall and Talbot's (68) work on neural mechanisms in vision. A major difference in their "statistical theory" appears to be in their emphasis on the role played by visual mechanisms, particularly the firing "on-off" receptor processes in the retina and also the process of physiological nystagmus. That is, even during fixation there are eye tremors reaching varying degrees of movement and taking in varying numbers

of cones. Gradients and peaks are established, and the cells mediating the "on-off" activity will become differently adapted by the rate and length of time of their excitation. A differential rate of recovery, which hinges on the time interval between inspection and test figures, is the crucial factor in the production of figural aftereffects. Osgood and Heyer conclude that, although their position does not constitute a disproof of Köhler's position, a firmer neurophysiological base brings it into the service of the law of parsimony.

Krauskopf (58) found a mean aftereffect displacement of 12.6′ when a test figure was presented 1.8 seconds after an inspection figure. The test period lasted only 0.3 second. Since the mean nystagmic fluctuation is said to be 4′, he questions the possibility that it could create aftereffects of the magnitude he achieved. Smith, who previously had raised six objections to the satiation theory (80), rephrased the same six as objections to Osgood and Heyer's statistical theory and added two more specific objections (81). Osgood (71) answered Smith's criticisms but had no solution for the apparent schism, pointed out by Smith, between functional and anatomical data. That is, how can the effects of an inspection figure, represented in one visual striate area, influence the perception of a test figure which is represented in the radically removed opposite striate area, across the median longitudinal fissure? Köhler and Wallach had previously suggested that the corpus callosum must be particularly conducive to satiational effects.

Smith later supported another of his objections by an experiment showing that test figures may move *toward* a previously satiated area (82). However, in two of his demonstrations at least, a part of the outline of the inspection figure coincides with a contour of the test figure. This coincidence occurs in the area of "attraction." Since it has frequently been demonstrated that coincidence of test and inspection figures causes no displacement, Smith's results may be explained on the basis of uneven testing of satiation, where the distance paradox is magnified and causes a flattening effect making the figures appear to approach each other, much as in the Walthall experiments (92). This type of explanation cannot be used in the third series of figures presented by Smith; and his demonstration of attraction in this third experiment, if confirmed, would present some interesting problems for any theory of figural aftereffects.

Hebb (37) also draws on Marshall and Talbot in criticizing satiation theory, particularly in relation to Köhler and Wallach's proposal that in peripheral vision perceived distances remain the same because of higher resistance. This permanent satiation is said to be caused by the squeezing of contours into the smaller cortical representation given the peripheral retina. According to Hebb, it is more probable that the frequency of contours which are subliminal for the peripheral field compensates for the greater spacing of contours in the macular area of the visual cortex.

SATIATION AND BRAIN DAMAGE

In this section there will be studies mentioned which on first glance may not seem relevant to the present review. For instance, some studies which utilize critical flicker fusion, autokinetic movement, stroboscopic movement, the rotating spiral, or other such apparatus will be reviewed even though many such studies make no mention of satiation theory. How-

ever, Köhler was originally led to figural aftereffects by way of stroboscopic movement; and from his reasoning it would seem to follow that, if satiation is a physiological fact, it must in some way enter into the perception of CFF, apparent motion, negative aftereffects of movement, and so on, even though temporal factors are somewhat different.

In 1911, Wohlgemuth (105) wrote a comprehensive paper on the aftereffect of seen movement and suggested the use of this technique as a diagnostic aid for such things as the study of the "optic tract," believing that aftereffects were peripheral rather than central. Modern perceptual laboratory techniques have been used extensively by Bender and Teuber and their associates on brain-damaged veterans. In their studies (5, 87) they found that with brain injury there was a decrease in CFF, a rise in threshold for apparent (stroboscopic) motion, and an abnormal increase of phenomenal (perceived) speed. Apparent displacements were increased by prolongation of exposure and decreased or abolished on tachistoscopic study. Some differences in effect were found with changes from monocular to binocular viewing.

Most relevant to the present review, however, is their attempt to relate their results on the brain-injured to Köhler and Wallach's results on satiated normals. They found the following points of agreement: the depth effect, associated with a pale, grayish appearance, was the most constant finding by both sets of researchers. Size effect, though less constant, was also found, along with distortion of lines and, at times, displacement of entire figures in the coronal plane. Temporary loss of the ability to use binocular parallax (on the stereo-

scope) was also noted. Finally, the tendency of all patients was to show more disturbances in the lower quadrants of the visual field.

Two other phenomena should be noted: autokinetic movement was away from the affected region of the field, and apparent displacements were increased by prolongation of exposure and minimized or abolished on tachistoscopic study.

The above results were based on patients who had penetrating injuries in an area crucial for visual examination, the occipital or parieto-occipital lobes. These patients were not first satiated and then tested. They performed as if they were satiated without prior stimulation. These facts are important to keep in mind in view of later research with brain injured subjects.

Köhler, referring to the work of Bender and Teuber, agrees that " . . . functional characteristics of injured areas seem to resemble those of strongly satiated areas. One feels tempted to assume that, in both cases, the impedance of the crucial regions is increased" (49, p. 242). Many of the following studies were designed essentially to test this possibility.

Kinesthetic Figural Aftereffects

Jaffe (43) tested the hypothesis that the cortical "electric fields" of patients with brain lesions would be altered and kinesthetic aftereffects consequently affected. His experimental subjects were war veterans with penetrating brain injuries and his apparatus consisted of strips of aluminum. The comparison strip was graduated in width from one to three inches. He found that 10 of his subjects with somatosensory defects showed so much variability when

the smallest strip was used as an inspection figure that the mean aftereffects with that strip did not reach significance. In all other tests, however, the group mean for amount and direction of kinesthetic aftereffect did not differ from the normals or from another group of 10 brain damaged patients without sensory defect. Although the brain lesions were unilateral, no significant differences were found between affected and unaffected hands.

These results do not support the hypotheses stemming from satiation theory, since Jaffe assumed that any patient with demonstrable sensory loss must have suffered direct damage to the cerebral somatosensory projections. They are also in direct contrast to the results of Klein and Krech (47) who used a very similar type of apparatus with a heterogeneous group of 12 severely brain injured subjects and measured not only the intensity of the kinesthetic figural aftereffect but also the rate of satiation and the persistence of the aftereffect. They reported that the aftereffects were more intense, were reached more rapidly, and persisted longer in the brain-injured than in a matched control group, and that the longer recovery rate was the single most sensitive diagnostic indicator. In fact, the tumor cases continued to build up, so that aftereffects were greater five minutes after a rest period than at any other point in the experiment.

This latter finding is most interesting in view of the fact that brain injured patients in general have a lowered threshold of CFF (1, 5) as well as difficulty in perceiving the negative aftereffect of the rotating Archimedes spiral (77). It is in line with the view that the brain-injured have

difficulty in turning back a force set in motion, whether this force is satiation or some other process.

In discussing the differences between his results and the results of Klein and Krech, Jaffe stresses the fact that both the locus of the lesion and the sensory state of the experimental subjects must be included in the data. Nevertheless, all we are told about his own patients is the laterality of damage and the fact that they are "roughly localized." Klein and Krech, on the other hand, explain their contradictory results as possibly being due to the greater organismic (i. e., metabolic) involvement stemming from extensive cortical damage, a view given some support by the findings of Wertheimer and his associates (101, 102). Local lesions, then, need not result in significant changes in properties of cortical conductivity. Klein and Krech found, in fact, that the more extensive the lesion, the more severely is satiation affected, thus accounting for Jaffe's negative results. They also discuss the possibility that satiation is a state of lowered cortical conductivity in a localized cortical area, and that an individual's satiability would indicate his basal level of cortical conductivity.

Although not strictly a measure of kinesthetic figural aftereffect, since there is no long inspection period, Weinstein's (93) work with brain injured patients seems to have significance at this point. His subjects were separated according to locus of lesion as well as presence or absence of somatosensory deficit. No matter which hand was used, the patients with parietal lobe lesions consistently overestimated the second of two identical weights to a greater degree than did the nonparietal brain-in-

jured or the controls. These results would seem to confirm the importance that the area of injury may have in certain types of sensory judgments.

Reversible, Ambiguous, and Hidden Figures

In 1939, Harrower (36) tested 30 patients, mainly postoperative cases with unilateral lesions. Although the sites of the lesions were indicated, the subjects were not separated into groups according to locus of damage. However, none had occipital lobe damage. Using the Rubin-Vase-Profile ambiguous figure, she found that compared with 30 controls the brain injured group showed a marked perseveration of the attitude engendered by their first perception. As a matter of fact, as many as 43% of the brain injured were unable to make anything out of the ambiguous figure. This rigidity does not by itself contradict satiation theory, however, since these patients were not damaged in the region crucial for visual satiation.

Yacorzynski and Davis (106) tested five patients with tumors extirpated from the frontal lobes and compared them with eight normal controls. The subjects fixated on a figure and reported changes in perspective. The experimental group reported fewer reversals than the control group, but unfortunately they were not compared with patients who had brain damage in nonfrontal areas. Incidentally, the experimental group also exaggerated the Müller-Lyer illusion.

Teuber, Battersby, and Bender (86) divided 60 war veterans with penetrating brain injuries into three groups on the basis of anterior, intermediate, and posterior penetration. The posterior group performed the poorest in almost all sections of a modified Gottschaldt hidden-figures test. Similarly, Teuber and Mishkin (89) found that patients with posterior brain lesions performed more poorly, when required to set a visual vertical against a conflicting visual background, than did a group with anterior brain lesions or a control group without brain injury.

On the other hand, Battersby and his associates (3), who also used a modified form of Gottschaldt's hidden figures, reported that although non-brain injured controls made fewer errors and took less time, brain damaged war veterans performed approximately the same regardless of the location of the cerebral lesion. However, it should be noted that their "posterior" group included patients with injuries in the "parieto-temperooccipital" areas. Similar findings have been reported by Teuber and Weinstein (90) who grouped their subjects according to locus and laterality of lesions and found that none of their groups differed one from the other in tests of overlapping contours, reversible figures, or hidden figures. Again the brain injured group as a whole differed significantly from a non–brain injured control group.

Thus we find that in two studies a deficit in figure-ground type tests was related to area of damage and in two other studies it was not. This figure-ground problem in organics, then, raises some interesting problems for satiation theory. In order for the theory to be consistent it must follow that patients with brain injuries in the occipital area would have difficulty in picking out hidden figures, since there should be a rapid rate of satiation and therefore a rapid shifting. With reversible or

ambiguous figures, once a percept is found, and with limited fixation, there should be a rapid reversal of figures. On the other hand, in patients injured in other than the occipital lobe or perhaps with general vascular type disorders, these results would not necessarily hold. In fact the opposite might be expected to occur, so that there would be some better capacity to pick out hidden figures but also greater rigidity and inability to reverse perspective. This kind of thinking could be further extended. For instance, it should also follow that patients with lesions in the visual areas would not see the Müller-Lyer illusion, or at least not see it as an illusion for very long, since Köhler and Fishback (52, 53) have reported that satiated normals are able to dispel the illusion. Naturally any results would have to be in the form of group averages, since large individual differences have been consistently reported. In order to more fully test these hypotheses, the research must be specifically designed. There must be, within the limits of possibility, clear evidence of the site of lesion and at least some fixation on the test object.

There are a few other studies which are very closely related to this problem. Werner and Strauss (96) found a disturbance in figure-ground differentiation in brain injured, mentally retarded children. The fact that these children were defective in visuo-motor and tactual-motor figure-ground tests would indicate the existence of this pathology in more than one sensory field. That the background becomes a distracting force is attested to by the work of Strauss and Lehtinen (83), and there is evidence of a similar type of disturbance in adult organics (104).

Apparent Movement and CFF

Werner and Thuma (97) reported a decrease in the threshold of CFF and a deficiency in the perception of apparent motion, but not in real motion, in brain-injured deficient children when compared with a matched endogenous group. While the indivijual variations of the brain injured group were large, and while most subjects of that group saw movement in a meaningful clock figure, only 2 out of 15 saw any movement in a simple line figure stroboscopically presented. Again we see the influence of the meaningfulness of test figures on the test results. Further, the threshold for simultaneity of the two successive stimuli was lower for the experimental group and there were greater differences in the CFF's of the organics when lower brightness intensities were used. The authors interpret their results as indicating defective integration and a perseveration of neural events which have been isolated by the brain lesions.

Despite the fact that the above findings were based on children with brain lesions acquired early in life, they are in general agreement with Bender and Teuber's (4, 5) findings based on an adult group with penetrating injuries acquired later in life. These authors found CFF thresholds lowered and stroboscopic thresholds raised, particularly in the affected quadrants.

Brenner (9) is highly critical of the results of Klein and Krech and of Shapiro's data. She eliminated the possibility of her subjects' being dependent on stimulus meaning, or of inferring movement, and then found that her six postencephalitic and six cerebral palsy patients paralleled nor-

mal subjects in their reports of the simultaneity, movement, and succession of tachistoscopically presented figures. In fact, at the upper thresholds the brain-damaged saw apparent movement at longer time intervals, a result opposite to what would be expected if satiation were considered an analogue of brain damage. It should be noted, however, that Klein and Krech specifically stated that their results applied to patients with severe brain injury, and that many of their subjects had extensive tumors or had experienced hemorrhages or severe head trauma, certainly a far different experimental group from Brenner's.

Bender and Teuber made use of the Plateau (Archimedes) spiral in their studies of the brain damaged and reported some interesting effects while the spiral was in motion. One of these effects was a reported fluctuation of the spiral. As for the negative aftereffect of motion, it was seen by their subjects both monocularly and binocularly and in any quadrant. Recent evidence has shown, however, that a great many brain damaged patients are unable to perceive the negative aftereffect of the spiral (28, 73, 77). These recent investigators are aware of the fact that many of the patients might have seen the aftereffect but not reported it, so a "testing the limits" phase has been added. The possibility certainly exists that a number of organic patients see the aftereffect but are unable to reconcile it with the fact that the machine is objectively stopped, and therefore they do not report it. It would certainly require an "abstract attitude" to do so.

Smith (80, 81) listed the negative aftereffect of the spiral as one of his evidences against the satiation and statistical theories of figural aftereffects, and although Osgood (71) replied that those theories were not required to cover all phenomena in the field of perception, Köhler had previously indicated that satiation might well account for such phenomena as stroboscopic movement. On the basis of the Osgood-Heyer statistical model, an attempt was made to answer Smith's criticism. According to George (30), the aftereffect is dependent upon the asymmetrical gradient between any two lines of the spiral. However, since after fixating on the expanding spiral the aftereffect of contraction may be observed by viewing any plain circular object, his explanation does not appear to be adequate. Perhaps a field theory of the spiral aftereffect will have to take the form of a satiation vector or movement, a directional spreading of impedance, somewhat analogous to the vector-field theory originally proposed by Brown and Voth (11).

Battersby (1, 2) tested normals and patients with frontal lobe and occipital lobe injuries on CFF three to six years after brain insult occurred. While the frontal lobe group showed no depression of CFF, those patients with occipital lobe injury and with hemianopsia had a significantly depressed threshold of CFF in the perimetrically normal half field. In the Columbia-Greystone studies (69, ch. 10) topectomized patients tested three weeks postoperatively showed a complete absence of autokinetic movement, but after three months they recovered the capacity to see such movement, although in simpler patterns. Thus it would seem that both locus of trauma and knowledge of time posttrauma are essential to certain perceptual studies of brain damaged subjects, possibly more

sensitively measured by apparent movement and CFF stimuli.

A perusal of the behavior of organics on certain tests, despite the fact that most such studies by the very nature of their experimental populations must suffer from a lack of control of the locus and extent of damage, brings out many inconsistencies in satiation theory. How is it that in brain damaged individuals the threshold of CFF is lowered while the threshold of stroboscopic movement is raised? Spatial distance is the major distinction between these two perceptual phenomena, but with heightened satiation this should make no difference.

Because of the many conflicting findings, it is difficult to see how satiation theory can adequately explain organic apperception. If we make the attempt we are led to the conclusion that most organics, who are extremely rigid, have a *low* capacity for satiation. We may perhaps add: except for those with injuries in the "crucial" area (i.e., occipital for vision, etc.). Although this nicely explains the organic's inability to shift, the lowered CFF threshold still needs to be explained. One answer might be that in normals continual flicker in one spot creates a highly satiated area which would impede the perception of the flickered light as continuous, thereby resulting in a high CFF threshold, while in stroboscopic movement a great deal of spread from the impeded area would enhance the perception of movement. In poorly satiable organics, impedance would be slow in building up and therefore their CFF thresholds would be generally lower and there would be a delay in the perception of stroboscopic movement. But we are still left with the conflicting results reported by

Klein and Krech, Jaffe, Wertheimer, and Brenner; and Klein and Krech's results would indicate *high* satiability no matter where the area of injury.

SATIATION, LEARNING, AND PERSONALITY THEORIES

Some of the more interesting extensions of the theory of satiation have been into the fields of learning, cognition, and personality. Köhler and Fishback (**52, 53**) used the phenomena resulting from fixations of the Müller-Lyer illusion as a springboard into the area of learning. Upon repeated presentations the illusion is erased, and in some cases a negative illusion is established. The substitution of evenly spaced vertical lines (test figures) after satiation creates an illusion consistent with satiation theory. The authors claim that the destruction of the Müller-Lyer illusion is a result of satiation, not practice effect in the usual sense. As evidence of this, they point out that illusions are destroyed without the subjects' knowing they are dealing with an illusion, that the process is not accelerated when the subjects are given this information, that a negative illusion can be established and that the effect is to a high degree localized. The fact that spaced trials are more effective than massed trials is said to be due to "over-satiation" in the latter instance and consequent deflection of currents into weaker areas, thus slowing the process.

But the authors go a step further. They imply a relationship between satiation and memory patterns and suggest that memory patterns may be mediated by the polarizing action of currents in the brain. As evidence of this possibility they cite the fact that some subjects who were tested three to six months later still showed a con-

siderable distortion. For ordinary memory, however, traces would have to be weak ones. They recognize the many problems involved in such a formulation, including the question of the temporal order of events and the nonlocalized condition of recognition and recall. In a similar vein, in a study of the auditory aftereffect, Deutsch (19) wondered if the practice effect might not in reality be due to prolonged stimulation rather than to practice or learning.

In a very interestingly designed research, Fox (26) conceived of the learning process as possessing striking similarities to the figure-ground phenomenon. In a problem-solving situation, for example, the tentative hypothesis emerges from the ambiguous ground and, failing to find closure, recedes and gives way to a new figure, a new hypothesis. When closure is achieved in the form of a solution, the field is articulated into a temporarily stable figure-ground relationship.

To test out this conception, Fox used 300 college students and compared the number of alternations perceived in the Rubin-Vase-Profile figures with the ability to solve 75 anagrams. As he predicted, there was a significant curvilinear relationship between the number of reversals and the number of correct solutions of anagrams. That is, a high degree of satiation as measured by a great amount of shifting on the reversible figure causes too much alternation of hypothesis to achieve the closure of any one hypothesis. A low degree of satiation causes too much rigidity. Fox admits the role that previous learning plays in his results and even concedes that the data could be explained by alternative hypotheses. Nevertheless, his predictions were based on satiation theory and his

findings are consistent with that theory. It should be noted that his finding of curvilinearity is very similar to Wertheimer's metabolic findings.

One implication of Fox's study is that an individual can be more "passive" and allow satiation to work for him, in a sense, and also more "active" when it is to his advantage. In other words, satiation is a process which is at the service of the entire organism. In this way one might explain how, after a rest period during which a particular problem is not primary, a sudden insight or solution to that problem will reach awareness. Halstead, in speaking of the process of abstraction, touches directly on this problem. "It is as if the cortex were able to throw the central object slightly out of focus and then bring it back into clear consciousness" (34, p. 266).

It would seem to the present reviewer that these kinds of extensions of the investigations of physiological perceptual processes into areas usually conceived of as more purely psychological are important moves and may more meaningfully provide us with tests of the generality of satiation theory. Few would doubt the importance of the roles played in perception both by physiological structure and by experience, but to understand these roles as fully as possible it would be well to know the nature and extent of individual differences in native endowment and physiological capacity. Perhaps we have here one means of doing just that in the area of perception, as well as in other areas.

Certainly the contributions of Eysenck (22) are significant in this respect. He cited evidence that causal factors which are responsible for extroverted behavior are to a large ex-

tent central in origin. From this evidence, and basing his thinking mainly on Hull's concept of reactive inhibition and to a somewhat lesser degree on Pavlov's work, he hypothesized that people who develop strong reactive inhibitions quickly and in whom dissipation is slow will exhibit extroverted patterns of behavior and, in case of nervous breakdown, hysterical disorders. With the reverse, we would find introversion and dysthymic disorders. Using kinesthetic figural aftereffects in the exact manner in which they were used by Klein and Krech, Eysenck's results were all in line with his predictions. As a group, the hysterics were less consistent than the dysthymics, and their figural aftereffects appeared more quickly, were stronger, and disappeared more slowly than did the aftereffects of the dysthymic group.

These results, of course, were group results. The aftereffects of some subjects were consistently in the opposite direction, a fact unaccounted for by any theory of figural aftereffects. In individual statistical tests, only the 30-second period of stimulation gave results significant at below the .05 level. Somewhat analogous to Klein and Krech's results, the differences between the groups were greater after the longer, 10-minute, rest period than after the shorter, 5-minute, rest. There was, in fact, a distinct similarity between Eysenck's hysterics and Klein and Krech's brain-damaged subjects on the kinesthetic aftereffect test. Klein's "self-inward" and "self-outward" groups also seem to be closely related to Eysenck's two groups. As Eysenck points out, however—in comparing the inhibition theory to the theory of cortical conductivity—in the former theory individuals are said to differ not with respect to conductivity but with respect to the rate at which inhibition is aroused by the passage of neural impulses.

The general line of approach taken by Shapiro and Eysenck, when compared with satiation theory, creates a curious marriage of Gestalt and Behavioristic concepts. As Eysenck notes, a description of reactive inhibition sounds very much like a description of satiation. As he further points out, there would also seem to be here a bridge between the field of learning and the field of perception.

Some time ago, McDougall (64) postulated a relationship between the rate of fluctuation of ambiguous figures and introversion-extroversion. He noted that certain drugs, which he called "intoxicating" or "extroverting" drugs, speeded up the rate of fluctuation, while "dream producing" or "introverting" drugs slowed down the rate of fluctuation. The picture would seem to be complete if we turn to the work on the effect of the metabolic rate on such perceptual phenomena as figural aftereffects and CFF, so that a relationship may be inferred between ability to satiate, metabolic rate, extroversion-introversion, and certain neurotic and psychotic disorders.

SUMMARY AND IMPLICATIONS

Ironically enough, the Gestalt emphasis on the whole organism does not fit with the point-to-point isomorphic aspects of satiation theory, itself questioned by cross-modal studies of figural aftereffects. Nor can the principles of closure and simplicity account for the destruction of circles apparent after fixation and "self-satiation," unless additional stipulations are added, as Luchins and Luchins point out.

It is quite evident from previous

considerations that although the theory of satiation very nicely explains figural aftereffects, it does less well when it moves into the field of ambiguous and hidden figures, apparent motion, and autokinesis, and breaks down or leads to absurdities when used as a complete explanation of the perception of the brain-damaged. This is not to say that satiation theory should be discarded at this time. It may very well be one of many physiological factors operating in perception. The failure would seem to be in the lengths to which some researchers go in order to fit observed phenomena into a specific theoretical framework.

There is room for further research in this area; and using the theory of satiation as a general frame of reference, despite some contradictory empirical findings, can still lead to fruitful experimentation. Sex differences in ability to satiate have been hinted at but not thoroughly explored. Can subjects be trained to satiate? Are there any differences in the speed of arousal and amount or persistence of the figural aftereffect as a function of age? This question takes on particular importance in view of the theoretical implication (63, 76) that homogeneity in the optical system is developed over a period of time by a process of statistical satiation. Is there any relationship between ability to satiate and intelligence? In this regard, the reviewer is currently engaged in research dealing with mental deficients.

The recognition by Fox (25) and by Sutherland (85) of multiple causative factors in the perception of figural aftereffects, and their ability to experimentally isolate and demonstrate these multiple factors, are important contributions and further investigations along these lines might be most productive. Other possible areas of future research have been touched on in the body of this review but one recent development should be mentioned at this point. Köhler and Wallach, in a discussion of possible future tests of satiation, make the following point:

... it will be remembered that in certain instances patients are not allowed to use their eyes for weeks on end. We do not know to what degree permanent satiation survives such a situation. If it is markedly lowered, cases of this kind would afterwards offer a chance of discovering whether with a reduced satiation of the whole field symmetrical objects exhibit any distortions which are normally absent (56, p. 346).

In point of fact, three observers returning from six days of wearing translucent plastic goggles which prevented pattern vision reported definite perceptual distortions. Among other things, size and position constancy were disturbed, straight edges and lines appeared to be curved, and in one subject there was a perceptual lag similar to that observed in brain damaged subjects. The distortions appeared most prominantly just outside the point of fixation (39).

Studies centering on satiation theory have contributed much to the field of perception. For one thing, many apparently isolated phenomena have been found to obey certain general laws. As an example, visual, kinesthetic, and auditory aftereffects seem to operate in the same general manner. Many hypotheses made on the basis of satiation theory have been borne out. Nevertheless, there is little doubt that another, more comprehensive theory will supersede, or more probably encompass, a modified theory of satiation. Whatever that theory may be, it will have to explain those consistent phenomena

now attributable to the process of satiation. In the meantime, however, satiation theory will have stimulated research, controversy, and discovery, which is all that can be asked of any theory.

REFERENCES

1. BATTERSBY, W. S. Critical flicker frequency in patients with cerebral lesions. *Amer. Psychologist*, 1950, **5**, 271–272. (Abstract)

2. BATTERSBY, W. S. The regional gradient of critical flicker frequency after frontal or occipital lobe injury. *J. exp. Psychol.*, 1951, **42**, 59–68.

3. BATTERSBY, W. S., KRIEGER, H. P., POLLACK, M., & BENDER, M. B. Figure-ground discrimination and the "abstract attitude" in patients with cerebral neoplasms. *A.M.A. Arch. Neurol. Psychiat.*, 1953, **70**, 703–712.

4. BENDER, M. B., & TEUBER, H.-L. Ring scotoma and tubular fields: their significance in cases of head injury. *A.M.A. Arch. Neurol. Psychiat.*, 1946, **56**, 300–326.

5. BENDER, M. B., & TEUBER, H.-L. Spatial organization of visual perception following injury to the brain. *A.M.A. Arch. Neurol. Psychiat.*, 1947, **58**, 721–739; 1948, **59**, 39–62.

6. BENDER, M. B., & TEUBER, H.-L. Disorders in visual perception following cerebral lesions. *J. Psychol.*, 1949, **28**, 223–233.

7. BEVAN, W., JR. The influence of figural after-effects upon visual intensity thresholds. *J. gen. Psychol.*, 1951, **45**, 189–207.

8. BRENNER, MAY W. Continuous stimulation and apparent movement. *Amer. J. Psychol.*, 1953, **66**, 494–495.

9. BRENNER, MAY W. The effect of brain damage on the perception of apparent movement. *J. Pers.*, 1956, **25**, 201–212.

10. BROWN, D. R. Investigation of the effects on cortical currents of organic damage to the brain by use of behavioral measures involving the kinesthetic figural after-effect. *Yearb. Amer. phil. Soc.*, 1954, 182–184. (Abstract)

11. BROWN, J. F., & VOTH, A. C. The path of seen movement as a function of the vector-field. *Amer. J. Psychol.*, 1937, **49**, 543–563.

12. BROWN, K. T. Methodology for studying figural after-effects and practice effects in the Müller-Lyer illusion. *Amer. J. Psychol.*, 1953, **66**, 629–634.

13. BROWN, K. T. Rate of apparent change in a dynamic ambiguous figure as a function of observation-time. *Amer. J. Psychol.*, 1955, **68**, 358–371.

14. CARLSON, V. R. Satiation in a reversible perspective figure. *J. exp. Psychol.*, 1953, **45**, 442–448.

15. CHYATTE, C. The relation of cortical potentials to perceptual functions. *Genet. Psychol. Monogr.*, 1954, **50**, 189–228.

16. CRUTCHFIELD, R. S., & EDWARDS, W. The effect of a fixated figure on autokinetic movement. *J. exp. Psychol.*, 1949, **39**, 561–568.

17. DAVIS, PAULINE A. Effects on the EEG of changing the blood sugar level. *A.M.A. Arch. Neurol. Psychiat.*, 1943, **49**, 186–194.

18. DEATHERAGE, B. H., & BITTERMAN, M. E. The effect of satiation on stroboscopic movement. *Amer. J. Psychol.*, 1952, **65**, 108–109.

19. DEUTSCH, J. A. A preliminary report on a new auditory after-effect. *Quart. J. exp. Psychol.*, 1951, **3**, 43–46.

20. EDWARDS, W., & CRUTCHFIELD, R. S. Differential reduction of autokinetic movement by a fixated figure. *J. exp. Psychol.*, 1951, **42**, 25–31.

21. ENGEL, G. L., & MARGOLIN, S. G. Clinical correlation of the electroencephalogram with carbohydrate metabolism. *A.M.A. Arch. Neurol. Psychiat.*, 1941, **45**, 890–892.

22. EYSENCK, H. J. Cortical inhibition, figural aftereffect, and theory of personality. *J. abnorm. soc. Psychol.*, 1955, **51**, 94–106.

23. FERNBERGER, S. W. A figural after-effect in the third dimension of visual space. *Amer. J. Psychol.*, 1948, **61**, 291–293.

24. FISICHELLI, V. R. Reversible perspective in Lissajous figures; some theoretical considerations. *Amer. J. Psychol.*, 1947, **60**, 240–249.

25. FOX, B. H. Figural after-effects: satiation and adaptation. *J. exp. Psychol.*, 1951, **42**, 317–326.

26. FOX, J. Köhler's satiation theory and individual differences in learning. Paper

read at West. Psychol. Ass., Fresno, April, 1952.

27. FREEBURNE, C. M., & HAMILTON, C. F. The effect of brightness on figural after-effect. *Amer. J. Psychol.*, 1949, 62, 567–569.

28. GALLESE, A. J., JR. Spiral aftereffect as a test of organic brain damage. *J. clin. Psychol.*, 1956, 12, 254–258.

29. GEORGE, F. H. On the figural after-effect. *Quart. J. exp. Psychol.*, 1953, 5, 128–135.

30. GEORGE, F. H. On the theory of the figural after-effect. *Canad. J. Psychol.*, 1953, 7, 167–171.

31. GIBSON, J. J. Adaptation, after-effect, and contrast in the perception of curved lines. *J. exp. Psychol.*, 1933, 16, 1–31.

32. GIBSON, J. J. Adaptation with negative after-effect. *Psychol. Rev.*, 1937, 44, 222–243.

33. GIBSON, J. J., & RADNER, M. Adaptation, after-effect, and contrast in the perception of tilted lines. I. Quantitative studies. *J. exp. Psychol.*, 1937, 20, 453–467.

34. HALSTEAD, W. C. Brain and intelligence. In L. A Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium.* New York: Wiley, 1951. Pp. 244–279.

35. HAMMER, ELAINE R. Temporal factors in figural after-effects. *Amer. J. Psychol.*, 1949, 62, 337–354.

36. HARROWER, M. R. Changes in figure and ground perception with cortical lesions. *Brit. J. Psychol.*, 1939, 30, 47–51.

37. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.

38. HELD, R. Patterning versus segregation in the production of visual after-effects. Paper read at East. Psychol. Ass., New York, April, 1957.

39. HERON, W., DOANE, B. K., & SCOTT, T. H. Visual disturbance after prolonged perceptual isolation. *Canad. J. Psychol.*, 1956, 10, 13–18.

40. HOCHBERG, J. E. Figure-ground reversal as a function of visual satiation. *J. exp. Psychol.*, 1950, 40, 682–686.

41. HOCHBERG, J. E., & BITTERMAN, M. E. Figural after-effects as a function of the retinal size of the inspection figure. *Amer. J. Psychol.*, 1951, 99–102.

42. HOCHBERG, J. E., & TRIEBAL, W. Figural after-effects with colored stimuli. *Amer. J. Psychol.*, 1955, 68, 133–135.

43. JAFFE, R. Kinesthetic after-effects following cerebral lesions. *Amer. J. Psychol.*, 1954, 67, 668–676.

44. JAFFE, R. The influence of visual stimulation on kinesthetic figural after-effects. *Amer. J. Psychol.*, 1956, 69, 70–75.

45. JONES, F. N., & BUNTING, ELIZABETH B. Displacement after-effects in auditory localization. *Amer. Psychologist*, 1949, 4, 389. (Abstract)

46. KLEIN, G. S. The personal world through perception. In R. R. Blake and G. V. Ramsey (Eds.), *Perception: An approach to personality.* New York: Ronald Press, 1951. Pp. 328–355.

47. KLEIN, G. S., & KRECH, D. Cortical conductivity in the brain injured. *J. Pers.*, 1952, 21, 118–148.

48. KÖHLER, W. *Dynamics in psychology.* New York: Liveright, 1940. Pp. 43–106.

49. KÖHLER, W. Relational determination in perception. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium.* New York: Wiley, 1951. Pp. 200–243.

50. KÖHLER, W., & DINNERSTEIN, DOROTHY. Figural after-effects in kinesthesis. In *Miscellanea psychologica Albert Michotte.* Paris: Librairie Philosophique, 1947. Pp. 196–220.

51. KÖHLER, W., & EMERY, D. A. Figural after-effects in the third dimension of visual space. *Amer. J. Psychol.*, 1947, 60, 159–201.

52. KÖHLER, W., & FISHBACK, JULIA. The destruction of the Müller-Lyer illusion in repeated trials: I. An examination of two theories. *J. exp. Psychol.*, 1950, 40, 267–281.

53. KÖHLER, W., & FISHBACK, JULIA. The destruction of the Müller-Lyer illusion in repeated trials: II. Satiation patterns and memory traces. *J. exp. Psychol.*, 1950, 40, 398–410.

54. KÖHLER, W., & HELD, R. The cortical correlate of pattern vision. *Science*, 1949, 110, 414–419.

55. KÖHLER, W., HELD, R., & O'CONNELL, D. N. An investigation of cortical currents. *Proc. Amer. phil. Soc.*, 1952, 96, 290–330.

56. KÖHLER, W., & WALLACH, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. phil. Soc.*, 1944, 88, 269–357.

57. KRAUSKOPF, J. Figural after-effects in auditory space. *Amer. J. Psychol.*, 1954, 67, 278–287.

58. KRAUSKOPF, J. The magnitude of figural after-effects as a function of the duration of the test period. *Amer. J. Psychol.*, 1954, **67**, 684–690.

59. LANDIS, C. Something about flicker-fusion. *Sci. Mon., N. Y.*, 1951, **73**, 308–314.

60. LASHLEY, K. S., CHOW, K. L., & SEMMES, J. An examination of the electrical field theory of cerebral integration. *Psychol. Rev.*, 1951, **58**, 123–136.

61. LIVSON, H. H. After-effects of prolonged inspection of apparent movement. *Amer. J. Psychol.*, 1953, **66**, 365–376.

62. LUCHINS, A. S., & LUCHINS, EDITH H. The satiation theory of figural after-effects and the Gestalt principles of perception. *J. gen. Psychol.*, 1953, **49**, 3–29.

63. LUCHINS, A. S., & LUCHINS, EDITH H. The satiation theory of figural after-effects and the principle of Prägnanz. *J. gen. Psychol.*, 1953, **49**, 185–199.

64. MCDOUGALL, W. *Outline of abnormal psychology.* New York: Scribner, 1926.

65. MARKS, M. R. A further investigation of the Köhler-effect. *Amer. J. Psychol.*, 1949, **62**, 62–74.

66. MARKS, M. R. Some phenomena attendant on long fixation. *Amer. J. Psychol.*, 1949, **62**, 392–398.

67. MARQUART, DOROTHY J. The satiational theory of figural after-effects. *J. gen. Psychol.*, 1954, **51**, 83–91.

68. MARSHALL, W. H., & TALBOT, S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. *Biol. Sympos.*, 1940, **7**, 117–164.

69. METTLER, F. A. (Ed.) *Selective partial ablation of the frontal cortex.* New York: Hoeber, 1949.

70. NACHMIAS, J. Figural after-effects in kinesthetic space. *Amer. J. Psychol.*, 1953, **66**, 609–612.

71. OSGOOD, C. E. Kendon Smith's comments on "A new interpretation of figural after-effects." *Psychol. Rev.*, 1953, **60**, 211–212.

72. OSGOOD, C. E., & HEYER, A. W., JR. A new interpretation of figural after-effects. *Psychol. Rev.*, 1952, **59**, 98–118.

73. PAGE, H. A., RAKITA, G., KAPLAN, H. K., & SMITH, N. B. Another application of the spiral aftereffect in the determination of brain damage. *J. consult. Psychol.*, 1957, **21**, 89–91.

74. PRENTICE, W. C. H. The relation of distance to the apparent size of figural after-effects. *Amer. J. Psychol.*, 1947, **60**, 617–627.

75. PRENTICE, W. C. H. On the size of the figural after-effect with varying distances. *Amer. J. Psychol.*, 1950, **63**, 589–593.

76. PRENTICE, W. C. H., & BEARDSLEE, D. C. Visual "normalization" near the vertical and horizontal. *J. exp. Psychol.*, 1950, **40**, 355–364.

77. PRICE, A. C., & DEABLER, H. L. Diagnosis of organicity by means of spiral aftereffect. *J. consult. Psychol.*, 1955, **19**, 299–302.

78. SHAPIRO, M. B. Experimental studies of a perceptual anomaly. III. The testing of an experimental hypothesis. *J. ment. Sci.*, 1953, **99**, 394–409.

79. SHAPIRO, M. B. A preliminary investigation of the effects of continuous stimulation on the perception of apparent movement. *Brit. J. Psychol.*, 1954, **45**, 58–67.

80. SMITH, K. R. The satiational theory of figural after-effects. *Psychol. Rev.*, 1948, **61**, 282–286.

81. SMITH, K. R. The statistical theory of figural after-effects. *Psychol. Rev.*, 1952, **59**, 401–402.

82. SMITH, K. R. "Atraction" in figural after-effects. *Amer. J. Psychol.*, 1954, **67**, 174–176.

83. STRAUSS, A. A., & LEHTINEN, LAURA E. *Psychopathology and education of the brain-injured child.* New York: Grune & Stratton, 1947.

84. SUMMERFIELD, A., & MILLER, K. M. Visual illusion and figural after-effect, with and without fixation. *Quart. J. exp. Psychol.*, 1955, **7**, 149–158.

85. SUTHERLAND, N. S. Figural after-effect, retinal size and apparent size. *Quart. J. exp. Psychol.*, 1954, **6**, 35–44.

86. TEUBER, H.-L., BATTERSBY, W. S., & BENDER, M. B. Performance of complex visual tasks after cerebral lesions. *J. nerv. ment. Dis.*, 1951, **114**, 413–429.

87. TEUBER, H.-L., & BENDER, M. B. Alterations in pattern vision following trauma of occipital lobes in man. *J. gen. Psychol.*, 1949, **40**, 37–57.

88. TEUBER, H.-L., & BENDER, M. B. Perception of apparent movement across acquired scotomata in the visual field. *Amer. Psychologist*, 1950, **5**, 271. (Abstract)

89. TEUBER, H.-L., & MISHKIN, M. Judgement of visual and postural vertical

after brain injury. *J. Psychol.*, 1954, **38**, 161–175.

90. Teuber, H.-L., & Weinstein, S. Ability to discover hidden figures after cerebral lesions. *A.M.A. Arch. Neurol. Psychiat.*, 1956, **76**, 369–379.

91. Walthall, W. J., Jr. The Köhler effect. *Amer. J. Psychol.*, 1946, **59**, 152–155.

92. Walthall, W. J., Jr. Field strength of the Köhler effect. *J. gen. Psychol.*, 1949, **41**, 27–32.

93. Weinstein, S. Time error in weight judgment after brain injury. *J. comp. physiol. Psychol.*, 1955, **48**, 203–207.

94. Weitz, J., & Compton, Bertita. A further stereoscopic study of figural after-effects. *Amer. J. Psychol.*, 1950, **63**, 78–83.

95. Weitz, J., & Post, Dorothy. A stereoscopic study of figural after-effects. *Amer. J. Psychol.*, 1948, **51**, 59–65.

96. Werner, H., & Strauss, A. A. Pathology of figure-background relation in the child. *J. abnorm. soc. Psychol.*, 1941, **36**, 236–248.

97. Werner, H., & Thuma, B. D. A deficiency in the perception of apparent motion in children with brain injury. *Amer. J. Psychol.*, 1942, **55**, 58–67.

98. Werner, H., & Thuma, B. D. Critical flicker-frequency in children with brain injury. *Amer. J. Psychol.*, 1942, **55**, 394–399.

99. Wertheimer, M. Constant errors in the measurement of figural after-ef-
fects. *Amer. J. Psychol.*, 1954, **67**, 543–546.

100. Wertheimer, M. The differential stability of schizophrenic and normal subjects: A test of a deduction from the theory of figural after-effects. *J. gen. Psychol.*, 1954, **51**, 291–299.

101. Wertheimer, M. Figural aftereffects as a measure of metabolic efficiency. *J. Pers.*, 1955, **24**, 56–73.

102. Wertheimer, M., Levine, H., & Wertheimer, Nancy. The effect of experimentally induced changes in metabolism on perceptual measures of metabolic efficiency. *Percept. Mot. Skills*, 1955, **5**, 173–176.

103. Wertheimer, M., & Wertheimer, Nancy. A metabolic interpretation of individual differences in figural aftereffects. *Psychol. Rev.*, 1954, **61**, 279–280.

104. Williams, H. L., Ardie, L., Gieseking, C., & Rubenstein, I. The relation of brain injury and visual perception to block design rotation test. *J. consult. Psychol.*, 1956, **20**, 275–280.

105. Wohlgemuth, A. On the after-effect of seen movement. *Brit. J. Psychol., Monogr. Suppl.*, 1911, **1**, 1–117.

106. Yacorzynski, G. K., & Davis, L. An experimental study of the functions of the frontal lobes in man. *Psychosom. Med.*, 1945, **7**, 97–107.

RESEARCH DEFINITIONS OF MENTAL HEALTH AND MENTAL ILLNESS[1]

WILLIAM A. SCOTT

University of Colorado

A serious obstacle to research in the area of mental illness lies in the lack of a clear definition of the phenomenon to be studied. The term "mental ill health" has been used by different researchers to refer to such diverse manifestations as schizophrenia, suicide, unhappiness, juvenile delinquency, and passive acceptance of an intolerable environment. Whether some or all of these various reactions should be included in a single category of "mental illness" is not clear from a survey of the current literature. Theories describing the nature and antecedents of one sort of disturbance rarely relate it to another, and there is a paucity of research evidence indicating the extent to which such manifestations are empirically intercorrelated.

In the face of such ambiguity it would appear useful to attempt an organized review of the various definitions of mental illness which are explicit or implicit in recent research, with a view toward highlighting their commonalities and discrepancies on both a theoretical and an empirical level. Such a presentation might help students concerned with causative factors to assess the comparability of previous research findings on correlates of "mental illness," and also point toward some next steps in research to discover the degree to which these diverse phenomena represent either unitary, or multifold, psychological processes.

The research criteria for mental illness to be reviewed here are subsumed under the following categories: (a) exposure to psychiatric treatment; (b) social maladjustment; (c) psychiatric diagnosis; (d) subjective unhappiness; (e) objective psychological symptoms; and (f) failure of positive adaptation. For each category we shall review studies which appear to have employed the definition, either explicitly or implicitly. This will be accompanied by a critical discussion of the adequacy of each definition, together with an assessment, based on empirical data where possible, of the relation between this and other definitions. Finally, we shall attempt to summarize the differences among the definitions, by indicating their divergent approaches to certain basic problems in the conceptualization of mental illness and health.

MENTAL ILLNESS AS EXPOSURE TO PSYCHIATRIC TREATMENT

The most frequently used operational definition of mental illness, at least in terms of the number of studies employing it, is simply the fact of a person's being under psychiatric treat-

ment. And this definition is usually restricted to hospital treatment, rather than outpatient service. Nearly all the ecological studies (e.g., **3, 16, 22, 30, 35, 50**) and most of the studies correlating mental illness with demographic characteristics (e.g., **5, 19, 29, 41, 47**) use this as a criterion. They obtain their information from hospital records or, in unusual instances (e.g., **28**), from psychiatrists in the area who furnish information about persons treated on an outpatient basis.

Such a definition of mental illness is operational rather than conceptual, but its implicit meaning for the interpretation of research results is that anyone who is regarded by someone (hospital authorities, relatives, neighbors, or himself) as disturbed enough to require hospitalization or outpatient treatment is mentally ill, and people who do not fit into such diagnoses are mentally healthy. Use of hospital records, moreover, requires that the criterion of the nature of the mental illness be the diagnosis which appears on the record.

Shortcomings of such an operational definition are recognized by no one better than its users. The reliability of psychiatric diagnosis is of course open to question, and any attempt to determine correlates of particular kinds of mental disturbance must take into account the large error inherent in the measuring process. (One study of the association between diagnosis at Boston Psychopathic Hospital and previous diagnoses of the patients at other hospitals showed only 51 per cent above-chance agreement between the two [cf. **15**, pp. 42–43].)

If "under the care of a psychiatrist" is to be regarded as the criterion of mental illness, one must realize the automatic limitation on the size of the mentally ill population that such a definition imposes. Kramer (**34**, p. 124) has estimated that the maximum possible number of mentally ill, under such a definition, would be less than 7,000,000, given the present number of available psychiatrists.

It has been suggested by both sociologists (**7, 10**) and physicians (**17**) that different rates of hospital admissions for different geographical areas may indicate more than anything else about the areas the relative degree to which the communities tolerate or reject persons with deviant behavior (**11**). Or as the Chief of the National Institute of Mental Health puts it: researchers using hospital records are dependent on the public's rather uneven willingness to give up its mentally ill members and to support them in institutions (**17**); this in addition to the admittedly unstandardized and often substandard methods of record-keeping used by the various hospitals is likely to render incomparable prevalence and incidence data from various geographical areas.

The effects of such differential thresholds for admission in various communities are difficult to estimate, since they cannot be uniform from study to study. In 1938 a house-to-house survey in Williamson County, Tennessee, yielded nearly one person diagnosed as psychotic, but never having been in a mental hospital, for every hospitalized psychotic from the county (**48**). By contrast, Eaton found in his study of the Hutterites (**14**) that more intensive canvassing by psychiatrists did not yield a larger number of persons deemed psychotic than did a more superficial count based on community reports.

Eaton's study *did* yield higher proportions of neurotic diagnoses the more intensive the case finding pro-

cedure became, and this observation relates to the finding in New Haven that neurotics under outpatient treatment came disproportionately from the upper socioeconomic strata (28). At first consideration, such differential rates seem readily attributable to the cost of psychiatric treatment, but Hollingshead and Redlich prefer to seek an explanation in the greater social distance between lower-class neurotics and the psychiatrists than in the case of middle- and upper-class neurotics. Whatever the sources of rate differences, it is clear that such correlations as have been reported make one wary of the hospital admissions or outpatient figures as indicative of the "true" incidence of psychiatric disorders. Thus the criterion of exposure to psychiatric treatment is at best a rough indicator of any underlying conceptual definition of mental illness.

MALADJUSTMENT AS MENTAL ILLNESS

Adjustment is necessarily determined with reference to norms of the total society or of some more restricted community within the society. Accordingly, one may conceptually define adjustment as adherence to social norms. Such a definition of mental health has an advantage over the preceding in encompassing a range of more-or-less healthy, more-or-less ill behavior, rather than posing a forced dichotomy. The operation for assessing mental health by this criterion might ideally be a community (or other relevant group) consensus concerning a given subject's degree of adjustment. This has been approximated by at least one set of studies (1, 2).

Rather than assess consensus by pooling many divergent individual opinions, it is possible to assume that a law or other visible sign of social norms constitutes the criterion against which adjustment is determined. Such reference is employed in studies of suicide (12, 26) or juvenile delinquency (25) or divorce (39, 53) as indicants of maladjustment. While the operational criterion may become dichotomous in such cases (whether or not the person comes in contact with the law), this is not necessarily so. Gordon (21) has suggested considering the "biologic gradient" of suicide, extending from contemplation of the act to its actual accomplishment.

Finally, it would be possible to assess degree of adjustment with reference to some externally defined set of requirements for a given social system. Thus a work situation might be seen as demanding a high level of productivity from all its members, and the degree of adherence to this standard becomes the criterion of adjustment, without reference to the individual opinions of the group members or to the manifest norms of the group. This criterion of conformity to the requirements of a given social structure has not been explicitly employed by any of the researchers covered in the present review, but it has been hinted at (37) and remains a possibility, provided that the structural requirements of a social system can be determined independently of the members' behaviors.

Theory of social structure suggests that these three criteria of adjustment would tend toward congruence: The demands of a particular social system lead to the development of social norms, which are expressed in laws or customs and also in the individual participants' notions of what is acceptable behavior. Lack of congruence may be taken as evidence of cultural lag, of poor correspondence

between manifest and latent function within the social structure, or of defensive psychological processes within the participating individuals. Since all of these factors supporting discrepancy do occur within most social systems, the criteria may be expected to yield somewhat different results.

When maladjustment is assessed by community consensus, one finds considerable divergence of opinion among various segments of the public regarding what constitutes good and poor adjustment. The Minnesota Child Welfare studies (1) showed differences in criteria for assessing adjustment among different occupational groups in the community. Teachers tended to emphasize standards different from those emphasized by ministers, who in turn displayed some differences from a more heterogeneous group of community adults. Beilin concludes that it is meaningless to discuss "adjustment" in the abstract or to contemplate the prediction of "adjustment" in general. One must specify *adjustment to what*, *adjustment to whose standards* (2). Lindemann reflects this relativistic conception of mental health when he states: "We find it preferable not to talk about a 'case' in psychiatry— rather we try to assess functional impairment in specific situations as viewed by different professional groups in the community. So a 'case' is really a relationship of possibly pathogenic situation and appropriate or inappropriate behavior to that situation. It is often a matter of arbitrary choice whether such a person becomes an object of psychiatric care" (38, p. 130).

Thus, though adjustment appears a more conceptually adequate criterion of mental health than does exposure to treatment, the necessity for considering different personal frames of reference and the demands of different social structures poses seemingly insurmountable obstacles to the establishment of mutually consistent operational definitions. All such difficulties which lie "hidden," as it were, under the psychiatric treatment criterion, come to the fore to plague the researcher trying to establish a criterion for adjustment which applies to the treated and nontreated alike.

PSYCHIATRIC DIAGNOSIS AS CRITERION FOR MENTAL ILLNESS

There have been a few studies in which entire communities or samples of them have been systematically screened, either by direct examination (44, 48) or by evidence from community records or hearsay (13, 14, 54). Here the criterion for mental illness or health need not be dichotomous, but can be divided into several gradations. Such intensive case-finding can be expected to increase the yield of persons classified as neurotic (34, p. 124) over that provided by the criterion of exposure to treatment, but whether the psychotic group is thereby increased will depend on the community (34, p. 124; 48) and, of course, on the standards for diagnosis employed by the particular investigator.

The lack of standardization of diagnostic procedures and criteria contributes to the incomparability of mental illness rates derived from such studies (34, p. 139; 55). So long as the criterion of assessment is largely dependent on the psychiatrist's subjective integration of a different set of facts for each subject, nonuniform results can be anticipated. Expensive and unreliable though the method may be, it at least places the judgment regarding

mental illness or health in the hands of professionals, which is not the case when adjustment is the criterion. And though hospitalization is in part determined by the judgment of professionals, *who* is sent to the hospitals for psychiatric diagnosis is, for the most part, out of the hands of the psychiatrists. As Felix and Bowers (17) have observed, it is the community rather than the clinician that operates the case-finding process today, and this will continue to be so until diagnostic examinations are given regularly to all people.

Mental Illness Defined Subjectively

It has been maintained by some that a major indication of need for psychotherapy is the person's own feeling of unhappiness or inadequacy. Conversely, the degree of mental health may be assessed by manifestations of subjective happiness, self-confidence, and morale. Lewis (36) quotes Ernest Jones to the effect that the main criterion for effect of therapy is the patient's subjective sense of strength, confidence, and well-being. Terman (52, 53) has used a "marriage happiness" test, composed largely of subjective items, and Pollak (43) has suggested that old-age adjustment be assessed in terms of the person's degree of happiness or well-being in various areas of his life.

That such criteria of mental health correlate somewhat with independent diagnoses by physicians has been indicated in two sorts of studies. In the Baltimore Eastern Health District (9), cases diagnosed psychoneurotic were found to express complaints about their own physical health; it is suggested that persons who report chronic nervousness can be classified as suffering from a psychiatric condi-

tion. Rogers has maintained that a marked discrepancy between one's "perceived self" and "ideal self" constitutes evidence of psychiatric disturbance (45), and some empirical studies lend support to this position. When Q sorts of subjects' self concepts are compared with Q sorts of their ideal selves, it is possible to distinguish psychiatric groups from non-psychiatric groups on the basis of the degree of discrepancy between these two measures (4). Furthermore, progress in therapy (as judged by the therapist) tends to be associated with increasing similarity between the patient's self concept and ideal self (46).

Though subjective well-being is an appealing criterion for mental health in ordinary daily living, it might be presumed that under some circumstances psychological defense mechanisms could operate to prevent the person's reporting, or becoming aware of, his own underlying unhappiness and disturbance. Jahoda (33) has rejected happiness as a criterion for mental health on somewhat different grounds: Happiness, she says, is a function not only of the person's behavior patterns, but of the environment in which he moves. If one wants to relate mental health to characteristics of the environment, then one must not take as a criterion of mental health something that already presupposes a benign environment. "There are certain circumstances in which to be happy would make it necessary first to be completely sick" (33, p. 105).

Such objections to this criterion imply that it is possible to find persons who are mentally ill by some other criterion, yet who nevertheless report themselves as happy or self-satisfied. Empirical demonstration of this implication is not available at present. In fact, while one study

predicted defensively high Q sorts for the self concept of paranoid psychotics, they were found to have a greater discrepancy between self- and ideal-sorts than normals, and no less discrepancy between these measures than psychoneurotics (4).

MENTAL ILLNESS DEFINED BY OBJECTIVE PSYCHOLOGICAL SYMPTOMS

It is generally accepted almost by definition that mental illness entails both a disordering of psychological processes and a deviation of behavior from social norms (6). The latter aspect of disturbance may be assessed as maladjustment to one's social environment (discussed above); the former aspect can presumably be assessed by psychological inventories aimed at the assumedly critical processes. The distinction between the psychological inventory approach and the subjective assessment procedure discussed above is not really a clear one. Subjective well-being may be regarded as one of the psychological processes which becomes disordered. Yet more "objective" measures of psychological process, which do not require the subject's verbal report of his degree of happiness, are frequently preferred, both to guard against purposeful distortion and to tap areas of disorder which may not be accompanied by subjective counterparts.

Such "objective" psychological inventories may represent various degrees of manifest purpose. For some, the objective of assessment is transparent, and the only reason they are not classed as devices for subjective report is that they stop just short of requiring the subject to report his over-all level of well-being. Such a manifest-level inventory is Halmos'

questionnaire concerning the respondent's difficulties in social relations (24).

At a somewhat less obvious level are such inventories as the MMPI, the War Department Neuropsychiatric Screening Battery, and the Cornell Medical Index, which require subjects to check the presence of various subjective and objective symptoms (e.g., "I smoke too much."). Once validated against an accepted criterion, such as psychiatric diagnosis, these are frequently used as criteria themselves. Rennie constructed a composite instrument of this type to assess his respondents' levels of mental health in the Yorkville study (44); at the same time, a validity analysis of the index was undertaken, by correlating each item with independent psychiatric diagnosis on a subsample of the respondents. On the basis of their experience with such a composite instrument, one of Rennie's colleagues (Langner, personal communication, August 1956) suggests caution in abstracting parts of previously validated batteries, since the item validities are sometimes not maintained when they are used out of context of the total instrument.

An adaptation of the psychiatric screening battery approach for use with children is suggested in the work of the St. Louis County Public Health Department (20). It involves obtaining information about symptoms from the children's mothers rather than from the children themselves. Naturally, the symptoms covered must be of the "objective" type ("Does Johnny wet the bed?") rather than of the "subjective" type ("Does Johnny worry a lot?"). As validated by an outside criterion (teachers' and psychiatric social workers' ratings of the child's level of adjustment), the

number of symptoms reported by the mothers appears to be a promising index of the child's mental health.

A general characteristic of the types of psychological inventories reviewed so far is that each item in the battery is assumed, a priori, to involve a "directional" quality, such that one type of answer (e.g., "yes" to "Are you troubled with nightmares?") may be taken as indicative of psychological disorder, and the opposite answer as indicative of normal functioning. Thus the index of disturbance is computed by adding all the positive indicators, weighted equally. That alternative methods of test construction may yield equally, or more, valid indices of mental illness is indicated by the extensive investigations of McQuitty (40).

McQuitty proposes several different methods of diagnostic test scoring, each based on explicit assumptions about the diagnostic procedure which the test is supposed to represent. One of the simplest assumptions, for example, is that an individual is mentally ill to the extent that his psychological processes deviate from the culturally modal processes. Thus, any type of multiple-alternative test may be administered to a group of subjects representing a "normal" population. Each alternative of each item is then scored for its "popularity." The score for a subject is then computed by adding the popularity scores of the items he checks (McQuitty calls this the T method of scoring); a high popularity score is taken as evidence of mental health (by this "typicality" criterion).

An alternative assumption proposed by McQuitty as underlying the diagnostic procedure might be that mental health is manifest to the degree that the subject's responses conform to any pattern of answers represented by a significant number of community people, regardless of whether that pattern is the most popular one. Such an assumption leads to a scoring procedure (H method) whereby a subject's index of "cultural harmony" is based on the degree to which his responses to different questions "go together" in the same manner as do the responses of all people in the sample who check the same alternatives he does.

Elaborations on these basic procedures provide for differential weighting of responses depending on their degree of deviance (WH method), and correction for "linkage" between successive pairs of items (WHc method).

The Bernreuter Personality Test and the Strong Vocational Interest Inventory were administered by McQuitty to a group of mental patients and to a group of university students; they were scored by different methods, the scores for the two tests were correlated, and the mean scores of the two groups compared. Results of the comparisons indicate that: (a) when appropriately scored, the Strong can discriminate mental patients from normals, though not so well as the Bernreuter; (b) better results are obtained if, instead of treating each answer as a separate, independent measure, it is evaluated in terms of the pattern of other answers with which it occurs (WHc scoring method); (c) within the Bernreuter, those items which correlated best with the total score (McQuitty's WHc method of scoring) and provided the best discrimination between patients and normals tended to be of the "subjective" type (i.e., they depended on the subject's intro-

spection, as in "Do you often have disturbing thoughts?") rather than the "objective" (items which an observer could report, such as "Do you talk very much?"); (*d*) different scoring procedures appeared differentially appropriate for the "subjective" and "objective" items; (*e*) when the "subjective" items were scored by the method most appropriate to them (i.e., the method which best discriminated patients from normals), and the "objective" items by their most appropriate method, the correlation between the two scores on the same group of subjects was about zero, indicating that two independent dimensions of mental health were being tapped by these two sets of items.

A separate study reported by McQuitty (40) indicated that the simple T method of scoring (based on the popularity of the subject's responses) both subjective and objective items significantly discriminated groups of school children classified on the basis of independent criteria of mental health. There is considerable evidence from these studies that, especially with respect to those traits measured by the "objective" items, the person may be regarded as mentally ill to the extent that he deviates from the dominant community pattern.

The foregoing studies provide a certain amount of evidence that measures of mental illness according to psychometric criteria relate to two of the criteria discussed earlier—maladjustment and psychiatric diagnosis. That such concurrent validation may yield somewhat different results from studies of predictive validity is indicated in Beilin's report of the Nobles County study (2). Two indices of student adjustment predictors were constructed, one (the "pupil in-

dex") based on students' responses to five different instruments, and the other (the "teacher index") based on teacher ratings. Both were concurrently validated against juvenile court judges' nominations of delinquent youngsters and against teachers' descriptions of the youngsters. Four years later the mental health of the youth was assessed by a number of different criteria—community reputation, interviewers' ratings, self-assessment, and an adaptation of the Rundquist-Sletto morale scale. The predictors correlated significantly with only some of the subsequent criteria, and all of the correlations were at best moderate. The "pupil index" correlated better with the interviewer's rating than with the community reputation criterion; while the "teacher index" correlated better with the subject's subsequent community reputation than with the interviewer's rating. Or, stated more generally, the psychologist's predictor predicted better to a psychologist's criterion, and a community predictor predicted better to a community criterion. Though the time span (four years) between the predictor and criterion measures may have been such as to allow for considerable change in the subjects, one is nevertheless reminded by these results that various criteria for mental health are not necessarily highly correlated.

In summarizing the various studies of mental health and illness defined by psychological testing batteries, we may note that many of them lack an underlying conception of the nature of mental illness from which to derive items and scoring procedures (a notable exception being McQuitty's measures), that some of them challenge the notion of the unidimen-

sional nature of mental health, and that their degree of correlation with other criteria, such as adjustment or psychiatric diagnosis, depends on the nature of the criterion.

MENTAL HEALTH AS POSITIVE STRIVING

A radically different approach to the assessment of mental health is indicated in the definitions proposed by some writers with a mental hygiene orientation. Gruenberg suggests that, though failure to live up to the expectations of those around him may constitute mental illness, one should also consider the person's failure to live up to his own potentialities (23, p. 131). Frank speaks of the "positive" aspect of mental health— healthy personalities are those who "continue to grow, develop, and mature through life, accepting responsibilities, finding fulfillments, without paying too high a cost personally or socially, as they participate in maintaining the social order and carrying on our culture" (18). In a less exhortative tone, Henry (27) discusses successful adaptation of the person in the "normal stressful situation." He sees many normal situations as situations of inherent stress. Some individuals in them develop mental disease, while others may develop out of them a more complex, but more successful, personality. It is this successful coping with the "normal stressful situation" that Henry regards as indicative of mental health.

Jahoda has translated this kind of emphasis on the positive, striving, aspects of behavior into a set of criteria amenable to empirical research. She proposes three basic features of mental health (31): (a) The person displays active adjustment, or attempts at mastery of his environment, in contrast to lack of adjustment or indiscriminate adjustment through passive acceptance of social conditions. (b) The person manifests unity of personality—the maintenance of a stable integration which remains intact in spite of the flexibility of behavior which derives from active adjustment. (c) The person perceives the world and himself correctly, independent of his personal needs.

Active mastery of the environment, according to Jahoda, presupposes a deliberate choice of what one does and does not conform to, and consists of the deliberate modification of environmental conditions. "In a society in which regimentation prevails, active adjustment will hardly be possible; in a society where overt regimentation is replaced by the invisible compulsiveness of conformity pressures, active adjustment will be equally rare. Only where there exists social recognition of alternative forms of behavior is there a chance for the individual to master his surroundings and attain mental health." (31, p. 563).

Such an approach is quite at odds with the subjective criterion of personal happiness, and with the conformity criterion referred to above as "adjustment." Attempted adjustment does not necessarily result in success, for success is dependent on the environment. The best mode of adjustment only maximizes the chances of success. It is mentally healthy behavior even if the environment does not permit a solution of the problem (33). Jahoda proposes that the criterion of happiness be replaced with some more "objective" definition of mental health, based on an explicit set of values.

In an unpublished community

study, Jahoda apparently attempted to assess only two of the aspects of mental health incorporated in her definition. Veridicality of perception (actually, of judgment) was determined by asking respondents to estimate certain characteristics of their communities concerning which objective data were available (e.g., proportion of people with only grade-school education), and at the same time inferring needs to distort reality from the respondent's evaluative statements about the problem (e.g., how important R believed education to be). This method of assessing need-free perception was regarded as something less than satisfactory (Jahoda, personal communication, August 1956), since the need was so difficult to determine, and it was difficult to establish unambiguously that distortion of judgment was due to the operation of a need rather than simply to lack of valid information.

The degree of attempted active adjustment was assessed by first asking a respondent to mention a particular problem in the community, then determining what he had done, or tried to do, about it, and how he felt about the problem at the time of interview (33). Three aspects of respondents' reactions were coded from their replies (32): (a) the stage of problem solution—mere consideration of the problem, consideration of solutions, or actual implementation; (b) the feeling tone associated with the problem—continued worry or improvement in feeling (either through partial solution or through passive acceptance); (c) the directness or indirectness of the approach—i.e., whether R went to the heart of the problem in his attempted solution or merely dealt temporarily with recurrent nuisances.

In her analysis Jahoda relates her measures of problem-solving and need-free perception to various characteristics of the respondents and of the communities in which they live. The relationships are interesting (e.g., in one of the communities the level of problem-solving was related to the degree of community participation of the respondent), but they appear to leave unanswered a basic question about the appropriateness of the criteria. If one accepts Jahoda's definition of mental health as involving the two components assessed in the study, then the results can be interpreted as showing what patterns of social interaction are associated with mental health. But if one is skeptical about the meaningfulness of the definition, then he is impelled to search for correlations between her two measures and other, more commonly accepted, criteria of mental health. These are not reported, although it would appear to be a fair question to ask about the relation of her concepts to those employed by other researchers.

If one is wedded to the happiness criterion of mental health, for example, one may speculate about the possibility of a negative relation between it and those provided by Jahoda. Unhappiness could conceivably lead to excessive coping behavior (attempted adjustment), or excessive coping behavior might elicit negative reactions from others which, in turn, would increase one's unhappiness. In like fashion, it could be that need-free perception would lead to increased unhappiness, since psychological defenses are not available to bolster one's self image. Though Jahoda might reject the suggestion that happiness is even relevant to her criteria, it would appear useful to ex-

plore, both conceptually and empirically, the interrelations among other measures of mental health and the novel one proposed by her.

Clausen (6) has maintained that researchers must ultimately face the task of relating mental health defined in positive terms to the individual's ability to resist mental illness under stress. At present it is not known whether they represent a common factor or are independent characteristics. Jahoda (personal communication, August 1956) suspects that positive mental health, as she defines it, may indeed represent a dimension orthogonal to that represented by the conventional psychological symptoms of mental illness. Thus, from a different approach than that employed by McQuitty comes the suggestion that mental health and illness may be a multidimensional phenomenon.

In employing these particular criteria, especially that of active adaptation, Jahoda seems willing to defend the evaluative standards implicit in it. And it may well be that values relating to attempted mastery of problems are every bit as defensible as the values of conformity implied in the adjustment criteria discussed above. Nevertheless, the former appear to exemplify the application of the Protestant ethic to the mental health movement in a manner which might introduce culture and class biases into one's conclusions. Miller and Swanson (42) have hypothesized that lower-class children will show more defeatism than middle-class children, as a result of different interpersonal and environmental experiences. Would they thereby be less mentally healthy by any standards besides those of the middle class? Truly, the problems posed in setting up absolute values from which to judge mental health and illness are perplexing.

BASIC PROBLEMS IN THE DEFINITION OF MENTAL HEALTH AND ILLNESS

Underlying the diversities in definition of mental illness one can discern certain basic differences of viewpoint concerning how the phenomena should be conceptualized. We may abstract certain foci of disagreement by posing the following four points of contention: (a) Does mental illness refer to a unitary concept or to an artificial grouping of basically different specific disorders? (b) Is mental illness an acute or chronic state of the organism? (c) Is maladjustment (or deviance from social norms) an essential concomitant of mental illness? (d) Should mental illness be explicitly defined according to values other than social conformity?

Each of the proposed definitions takes a stand, either explicitly or implicitly, on one or more of these issues. It is likely that resolution of disagreements will depend in part on the outcome of future empirical research. But at least some of the divergence inheres in the theoretical formulation of the problem, and is more a matter of conceptual predilection than of empirical fact. In either case, if one is to arrive at consistent theoretical and operational definitions of mental illness, it would be well to make explicit one's bias concerning each of these issues, and attempt to rationalize it in terms of his conception of the causes of disturbance.

THE UNITARY OR SPECIFIC NATURE OF MENTAL ILLNESS

The position that mental illness is manifest in some rather general form,

regardless of the specific diagnostic category in which the patient is placed, would appear to be implicit in the subjective definition of the phenomenon. If the person's feeling of happiness or adequacy is regarded as the crucial indicator of his mental state, this would appear to imply that over-all health or illness can be assessed for a particular person, regardless of the area of functioning referred to. Likewise, the definition of mental health in terms of purposeful striving or active adjustment tends to ignore differences in the underlying bases for such striving or lack thereof. Such a position has been stated explicitly by Stieglitz: "The mensuration of health . . . closely parallels the measurement of biological age as contrasted to chronological age . . . . We are no longer seeking to discover specific disease entities, or even clinical syndromes, but attempting to measure biological effectiveness in adaptation" (51, p. 79). And such a unitary view of the phenomenon is implied in Schneider's comment: "The major 'cause' of mental disease is seen as some form of disorientation between the personality and society" (49, p. 31).

By contrast, the specific view of mental illness is taken by Gordon: "What we choose to call mental disease is an artificial grouping of many morbid processes. The first essential, in my opinion, is to separate the various entities, and in the approach to an epidemiology of mental diseases, to center attention on some one condition, or a few selected conditions, which have functions in common with other mass diseases well understood in their group relationships" (15, p. 107). McQuitty offers empirical evidence in favor of a specific view, in his isolation of two quite independent measures of mental illness (by psychological testing), both of which

correlate with external diagnostic criteria. And he further speculates that the number of areas in which the degree of personality integration varies rather independently is probably greater than the two which he has isolated. "One might expect that mental illness might develop within any one or more patterns. In order to understand the mental illness of a particular subject, we must isolate the pattern, or patterns, of characteristics to which his mental illness pertains" (40, p. 22).

While the weight of opinion and evidence appears to favor the multidimensional view, this may simply be a function of the operational definitions employed (e.g., mental health defined by responses to a battery of tests is bound to turn out multidimensional to the extent that intercorrelations among the test items are low). But there are yet insufficient empirical data collected from the unitary point of view to test whether its assumption is correct. Indeed, it seems quite plausible that both happiness and active adaptation may be partially a function of the situation, hence the concept of mental health implied by them must become multidimensional to the extent that they allow for intersituational variability.

THE ACUTE OR CHRONIC NATURE OF MENTAL ILLNESS

The psychologist's testing approach to assessing mental illness inclines him toward a view of the condition as chronic. That is, the predisposing conditions within the organism are generally presumed to be relatively enduring, though perhaps triggered off into an actual psychotic break by excessively stressful situations. The epidemiological approach, on the other hand, is usually concerned with the counting of actual hospitalized cases, and this may incline one

toward a view of mental illness as predominantly acute. Felix has espoused this position explicitly: "Unless the kinds of mental illness are specified, I can't conceive that mental illness is a chronic disease. More mental illnesses by far are acute and even short term than there are mental illnesses which are chronic and long term." (15, p. 163). Of course, the epidemiological approach traditionally considers characteristics of the host, as well as characteristics of the agent and the environment. But the predisposing factors within the organism seem to be regarded, like "low resistance," not as a subliminal state of the disease, but rather as a general susceptibility to any acute attack precipitated by external factors.

It is easier to regard a psychosis as acute than it is similarly to regard a neurosis, since in the former disorder the break with normal behavior appears more precipitate. However, such a judgment, based on easily observable external behaviors, may be unduly superficial. Even in the case of such a discrete disturbance as suicide, at least one writer (21) recommends considering the biologic gradient of the disorder. He distinguishes varying degrees of suicide, with successful accomplishment as merely a possible end product. Where such continuity between morbid and nonmorbid states can be discerned, the possibility of chronic disturbance might well be considered.

THE PROBLEM OF MENTAL HEALTH AS CONFORMITY TO SOCIAL NORMS

The criterion of mental health based on adjustment clearly implies that conformity to the social situation in which the individual is permanently imbedded is a healthy response. And such an assumption would appear to be lurking, in various shapes, behind nearly all of the other definitions considered (with the possible exception of some of the "positive striving" criteria, which stress conformity to a set of standards independent of the person's immediate social group). In fact, McQuitty's methods of scoring psychological inventories are all explicitly based on the assumption that conformity (either to the total community or to a significant subgroup) is healthy.

If the stability of the larger social system be regarded as the final good, or if human development be seen as demanding harmony in relation to that social system, then such an assumption would appear basic and defensible. But one is still impelled to consider the possibility that the social system, or even an entire society, may be sick, and conformity to its norms would constitute mental illness, in some more absolute sense. If any particular behavior pattern is considered both from the standpoint of its adaptability within the social structure to which the individual maintains primary allegiance and from the standpoint of its relation to certain external ideal standards imposed by the observer, perhaps a comparison of the two discrepancy measures would yield information about the degree to which the social system approaches the ideal. On the other hand, such a comparison might be interpreted as merely indicating the degree to which the researcher who sets the external standards is himself adapted to the social system which he is studying. The dilemma appears insoluble.

THE PROBLEM OF VALUES IN CRITERIA FOR MENTAL HEALTH

The mental hygiene movement has traditionally been identified with one or another set of values—ideal standards from which behavior could be assessed as appropriate or inappropri-

ate.   The particular set of values adopted probably depends to a considerable degree on who is doing the judging.   Such a diversity of evaluative judgments leads to chaos in the popular literature and to considerable confusion in the usage of the term "mental health" in scientific research.   Kingsley Davis (8) presented a rather strong case for the proposition that mental hygiene, being a social movement and source of advice concerning personal conduct, has inevitably been influenced by the Protestant ethic inherent in our culture.   The main features of this Protestant ethic, as seen by him, are its democratic, worldly, ascetic, individualistic, rationalistic, and utilitarian orientations.

To the extent that research on mental health is based on criteria devolved from such an ideology, it is middle-class-Protestant biased.   To the extent that it is based on some other set of "absolute" norms for behavior, it is probably biased toward some other cultural configuration. At least one researcher, Jahoda (33), has clearly taken the position that mental health criteria must be based on an explicit set of values. There is some advantage in allowing the assumptions to come into full view, but in this case the resulting criteria appear to be rather specialized and not comparable with those used by other researchers.   Perhaps the difficulty lies not so much in the existence of explicit assumptions as in their level of generality. If a more basic set of assumptions could be found, from which the diverse criteria for mental health and illness can be derived, then comparability among researches might better be achieved. One would be in a better position to state when mental illness, as defined by psychological tests or by absence of active adjustment, is likely to be displayed in mental illness defined by psychiatric diagnosis or deviance from community standards.

Summary

The various categories of definitions of mental illness discussed here have been distinguished primarily on the basis of their differing operational definitions: the dependent variables employed in empirical research on the phenomena are clearly different. Moreover the conceptualizations of mental illness explicit or implicit in the empirical criteria are often quite divergent—viz., the radically different viewpoints underlying the "maladjustment," "subjective unhappiness," and "lack of positive striving" definitions.

Certain conceptual and methodological difficulties in each of these types of definition have been noted: "Exposure to treatment" is deficient in that only a limited proportion of those diagnosable as mentally ill ever reach psychiatric treatment. "Social maladjustment" is open to question because of the varying requirements of different social systems and the diversity of criteria for adjustment employed by community members. "Psychiatric diagnosis" provides an expensive, and often unreliable, method of assessing the state of mental health. "Subjective unhappiness" can be criticized as a criterion since it may be a function of intolerable environmental conditions as well as the psychological state of the person, and is subject to distortion by defense mechanisms. The validity of "objective testing procedures" appears to depend considerably on the method by which they are scored, and there is strong evidence that a major component of their score may simply be the

degree of conformity of the person to the community average. Finally, criteria included under the heading of "positive striving" are subject to question in that they are inevitably based on disputable value systems of their proponents.

While many of these difficulties would not be considered damaging from the point of view of certain of the definitions of mental illness, they run into conflict with others. Also they suggest certain basic incompatibilities among the various approaches to conceptualization of mental illness. Whether these incompatibilities should be reconciled by further theoretical and empirical exploration, or whether they should be regarded as valid indicators that mental health and illness constitute multidimensional phenomena is still a moot question. We can only note that various studies employing two or more of these different categories of criteria have tended to yield moderate, but not impressive, interrelations.

The criterion of "exposure to psychiatric treatment" has been related to "maladjustment," "psychiatric diagnosis," "subjective unhappiness," and "objective psychometrics." Also "maladjustment" has been related to "psychiatric diagnosis" and to certain "objective" measures; and "psychiatric diagnosis" has been related to both "subjective" and "objective" measures of mental illness. The areas of interrelationship for which no empirical studies have been found are between "subjective" measures and both "maladjustment" and "objective" assessment; also between the "positive striving" criteria and all of the other types of measures.

Two directions for future theory and research are indicated by these results. First, more investigations are needed of the extent of relationship among the various criteria, and of the conditions under which the magnitudes of the intercorrelations vary. Second, assuming absence of high intercorrelations under many conditions, it would be worthwhile to explore the implications of poor congruence between one measure and another—implications both for the person and for the social system in which he lives.

REFERENCES

1. BEILIN, H. The effects of social (occupational) role and age upon the criteria of mental health. *J. soc. Psychol.*, in press.
2. BEILIN, H. The prediction of adjustment over a four year interval. *J. clin. Psychol.*, 1957, **13**, 270–274.
3. BELKNAP, I. V., & JACO, E. G. The epidemiology of mental disorders in a political-type city, 1946–1952. In *Interrelations between the social environment and psychiatric disorders*. N. Y.: Milbank Memorial Fund, 1953.
4. CHASE, P. Concepts of self and concepts of others in adjusted and maladjusted hospital patients. Unpublished doctor's dissertation, Univer. of Colorado, 1956.
5. CLARK, R. E. Psychoses, income and occupational prestige. *Amer. J. Sociol.*, 1949, **54**, 433–440.
6. CLAUSEN, J. A. *Sociology and the field of mental health*. N. Y.: Russell Sage Foundation, 1956.
7. CLAUSEN, J. A., & KOHN, M. L. The ecological approach in social psychiatry. *Amer. J. Sociol.*, 1954, **60**, 140–151.
8. DAVIS, K. Mental hygiene and the class structure. *Psychiatry*, 1938, 1, 55–65.
9. DOWNES, JEAN, & SIMON, KATHERINE. Characteristics of psychoneurotic patients and their families as revealed in a general morbidity study. *Milbank Memorial Fund Quarterly*, 1954, **32**, 42–64.
10. DUNHAM, H. W. Current status of ecological research in mental disorder. *Social Forces*, 1947, **25**, 321–326.

11. DUNHAM, H. W. Some persistent problems in the epidemiology of mental disorders. *Amer. J. Psychiat.*, 1953, **109**, 567–575.

12. DURKHEIM, E. *Le suicide.* Paris: F. Alcan, 1897. (English translation, Glencoe, Ill.: Free Press, 1951.)

13. EATON, J. W. *Culture and mental disorders.* Glencoe, Ill.: Free Press, 1955.

14. EATON, J. W., & WEIL, R. J. The mental health of the Hutterites. In A. M. Rose (Ed.), *Mental health and mental disorder.* N. Y.: Norton, 1955.

15. *Epidemiology of mental disorder.* N. Y.: Milbank Memorial Fund, 1950.

16. FARIS, R. E. L., & DUNHAM, H. W. *Mental disorders in urban areas.* Chicago: Chicago Univer. Press, 1939.

17. FELIX, R. H., & BOWERS, R. V. Mental hygiene and socio-environmental factors. *Milbank Memorial Fund Quarterly*, 1948, **26**, 125–147.

18. FRANK, L. K. The promotion of mental health. *Ann. Amer. Acad. of Pol. Soc. Sci.*, 1953, **286**, 167–174.

19. FRUMKIN, R. M. Occupation and major mental disorders. In A. M. Rose (Ed.), *Mental health and mental disorder.* N. Y.: Norton, 1955.

20. GLIDEWELL, J. C., ET AL. Behavior symptoms in children and degree of sickness. *Amer. J. Psychiat.*, 1957, **114**, 47–53.

21. GORDON, J. E., ET AL. An epidemiologic analysis of suicide. In *Epidemiology of mental disorder.* N. Y.: Milbank Memorial Fund, 1950.

22. GRUENBERG, E. M. Community conditions and psychoses of the elderly. *Amer. J. Psychiat.*, 1954, **110**, 888–896.

23. GRUENBERG, E. M. Comment in *Interrelations between the social environment and psychiatric disorders.* N.Y.: Milbank Memorial Fund, 1953.

24. HALMOS, P. *Solitude and privacy.* London: Routledge and Kegan Paul, 1952.

25. HATHAWAY, S. R., & MONACHESI, E. D. The Minnesota Multiphasic Personality Inventory in the study of juvenile delinquents. In A. M. Rose (Ed.), *Mental health and mental disorder.* N. Y.: Norton, 1955.

26. HENRY, A. F., & SHORT, J. *Suicide and homicide.* Glencoe, Ill.: Free Press, 1954.

27. HENRY, W. E. Psychology. In *Interrelations between the social environment and psychiatric disorders.* N. Y. Milbank Memorial Fund, 1953.

28. HOLLINGSHEAD, A. B., & REDLICH, F. C. Social stratification and psychiatric disorders. *Amer. sociol. Rev.*, 1953, **18**, 163–169.

29. HYDE, P. W., & KINGSLEY, L. V. Studies in medical sociology. I: The relation of mental disorders to the community socio-economic level. *New England J. Med.*, 1944, **231**, 543–548.

30. JACO, E. G. The social isolation hypothesis and schizophrenia. *Amer. sociol. Rev.*, 1954, **19**, 567–577.

31. JAHODA, MARIE. Toward a social psychology of mental health. In A. M. Rose (Ed.), *Mental health and mental disorder.* N. Y.: Norton, 1955.

32. JAHODA, MARIE. The meaning of psychological health. *Soc. Casewk*, 1953, **34**, 349–354.

33. JAHODA, MARIE. Social psychology. In *Interrelations between the social environment and psychiatric disorders.* N. Y.: Milbank Memorial Fund, 1953.

34. KRAMER, M. Comment in *Interrelations between the social environment and psychiatric disorders.* N. Y.: Milbank Memorial Fund, 1953.

35. LEMERT, E. M. An exploratory study of mental disorders in a rural problem area. *Rural Sociol.*, 1948, **13**, 48–64.

36. LEWIS, A. Social aspects of psychiatry. *Edinburgh med. J.*, 1951, **58**, 241–247.

37. LINDEMANN, E., ET AL. Minor disorders. In *Epidemiology of mental disorders.* N. Y.: Milbank Memorial Fund, 1950.

38. LINDEMANN, E. Comment in *Interrelations between the social environment and psychiatric disorders.* N. Y.: Milbank Memorial Fund, 1953.

39. LOCKE, H. *Predicting adjustment in marriage: a comparison of a divorced and a happily married group.* N. Y.: Holt, 1951.

40. McQUITTY, L. L. Theories and methods in some objective assessments of psychological well-being. *Psychol Monogr*, 1954, **68**, No. 14.

41. MALZBERG, B. *Social and biological aspects of mental disease.* Utica: State Hosp. Press, 1940.

42. MILLER, D. R., & SWANSON, G. E. A proposed study of the learning of techniques for resolving conflicts of impulses. In *Interrelations between the social environment and psychiatric disorders.* N. Y.: Milbank Memorial Fund, 1953.

43. POLLAK, O. Social adjustment in old age. *Soc. Sci. Res. Council Bull.* No. 59, 1948.

44. RENNIE, T. A. C. The Yorkville com-

munity mental health research study. In *Interrelations between the social environment and psychiatric disorders.* N. Y.: Milbank Memorial Fund, 1953.

45. ROGERS, C. *Client-centered therapy.* Boston: Houghton Mifflin, 1951.

46. ROGERS, C., & DYMOND, ROSALIND. *Psychotherapy and personality change.* Chicago: Univer. of Chicago Press, 1954.

47. ROSE, A. M., & STUB, H. R. Summary of studies on the incidence of mental disorders. In A. M. Rose (Ed.), *Mental health and mental disorder.* N. Y.: Norton, 1955.

48. ROTH, W. F., & LUTON, F. H. The mental health program in Tennessee. *Amer. J. Psychiat.,* 1943, 99, 662–675.

49. SCHNEIDER, E. V. Sociological concepts and psychiatric research. In *Interrelations between the social environment and psychiatric disorders.* N. Y.: Milbank Memorial Fund, 1953.

50. SCHROEDER, C. W. Mental disorders in cities. *Amer. J. Sociol.,* 1942, 48, 40–47.

51. STIEGLITZ, E. J. The integration of clinical and social medicine. In I. Galdston (Ed.), *Social medicine—its derivations and objectives.* N. Y. Acad. of Med., 1947. N. Y.: Commonwealth Fund, 1949.

52. TERMAN, L. M., ET AL. *Psychological factors in marital happiness.* N. Y.: McGraw-Hill, 1938.

53. TERMAN, L. M., & WALLIN, P. The validity of marriage prediction and marital adjustment tests. *Amer. sociol. Rev.,* 1949, 14, 497–505.

54. TIETZE, C., ET AL. Personal disorder and spatial mobility. *Amer. J. Sociol.,* 1942, 48, 29–39.

55. TIETZE, C., ET AL. A survey of statistical studies on the prevalence and incidence of mental disorders in sample populations. *Publ. Hlth Rep.,* 1943, 58, 1909–1927.

PARAMETERS RELEVANT TO DETERMINING THE EFFECT OF EARLY EXPERIENCE UPON THE ADULT BEHAVIOR OF ANIMALS[1]

JOHN A. KING[2]

Roscoe B. Jackson Memorial Laboratory, Bar Harbor, Maine

The hypothesis that some experiences early in the life of an organism profoundly and persistently affect the behavior of the adult has recently been subjected to a large number of experimental tests with animal subjects. Some investigators (**15, 19, 42**) have approached this general hypothesis from the viewpoint of Freudian theory, which suggests that the development of an infant is arrested or fixated at an early age (**11**). Other investigators (**8, 9, 20**) have oriented their experiments in terms of Hebbian theory, which stresses the importance of early perceptual experiences on later performance in a learning situation (**16**). The approach of other investigators (**24, 43**) has been to test for a critical period in which a given experience has a greater effect on adult behavior than the same experience at a later age (**41**). Regardless of the approach, all of the experiments reviewed here have attempted to test whether some early experiences have a profound and lasting effect on later behavior.

Since the test of this hypothesis ultimately depends upon the adequacy of the methods used, it is desirable to examine these methods. The isolation of variables in any "longitudinal approach to the study of animal life is obviously more beset with methodological problems than studies concerned with only one portion of the life span" (**2**, p. 259). Although all the variables affecting development cannot be isolated and, perhaps, recognized, it is the purpose of this paper to analyze the recognizable parameters in the study of the effects of early experience on later behavior. This analysis deals chiefly with methodology, rather than results and conclusions. Most of the material is drawn from experimental studies on nonhuman mammals, particularly the rat and mouse, which show comparable rates of maturation.

Seven recognizable variables are of particular significance to the experiments under examination. They are: (*a*) the age of the animal when the experience is given, (*b*) age at the time of the test, (*c*) the duration or quantity of the experience, (*d*) the type or quality of the experience, (*e*) the type of the performance task required of the adult animal, (*f*) the method for testing persistence of the effects, and (*g*) the relation of the experience to the genetic background of the animal.

AGE AT THE TIME OF THE EXPERIENCE

In considering the effect of time in relation to the commonly accepted theory that the past experience of an animal determines in part his present behavior, it is possible to make three alternate assumptions: (*a*) that the same

experience at any time of life will have the same effect, (*b*) that the same experience will have a greater effect early in development and in proportion to the youth of the subject, and (*c*) that there are certain critical periods in development (some of which may occur early in life), during which the same experience may have a much more profound effect than at other times.

This quotation from Scott and Marston (**41**, p. 26) is followed by a refutation of the first two possibilities and a confirmation of the last, drawn largely from embryological and observational evidence. If the critical period hypothesis is to be adequately tested, the age of the animal at the time it is given the experience becomes particularly important. A test of the hypothesis demands answers to the questions: when do these critical periods occur; and, if they occur early in life, what is early? For purposes of age comparisons, a number of investigations on the rat and mouse are illustrated in Fig. 1. The age when the mouse or rat has been exposed to a given "early" experience varies from birth until over 100 days of age. Certainly, in view of this range, "early" does not mean the same thing to all investigators. Upon examination of Fig. 1, three age patterns appear in reference to the maturational processes of the rat or mouse.

First, there are *infantile* studies in which the experimental treatment is administered during the period from birth until weaning in the rat. Some investigators have treated their subjects over the entire 20- or 25-day period (**14, 26, 28, 42**), while others have selected a shorter interval during this period (**15, 43**). The second age period can be called the *juvenile* period, extending from weaning to sexual maturity—a period of approximately 40 days in the rat or mouse. Again, the entire period may be used

for treating the rat (**4, 50**) or shorter periods of 10 to 15 days (**19, 29**). The third period of treatment begins sometime before *sexual maturity* and continues into the adulthood of the rat (**1, 8, 9, 20**).

This classification of developmental periods of the rat into infantile, juvenile, and adult utilizes only the conspicuous changes induced by birth, weaning, and sexual maturity. These convenient landmarks are quite arbitrarily chosen, for subtle neurological or physiological modifications may actually be more important to the hypothesis under examination. However, in many investigations even these developmental changes of weaning and sexual maturity are not distinguished in administrating the treatment. Although this general approach to the study of the age variable is acceptable in the absence of physiological and neurological correlates to behavioral development, systematic examination of the different age periods is necessary. In experiments not designed to study the age variable, the treatment period should terminate before maturity, while most maturational processes are active. This procedure will distinguish experiments on the effects of early experience from those not directly concerned with maturation.

AGE AT THE TIME OF THE TEST

The age variable must be controlled during the testing period as well as during the treatment period. The crosses in Fig. 1 illustrate the age when the pretest or test period began in the experiments represented. Most experimenters began the testing procedures after the rats were full-grown. The exceptions to this practice have been studies in which weight (**3, 32, 42, 50, 51**), temperatures (**50**), or behavioral changes (**37**,
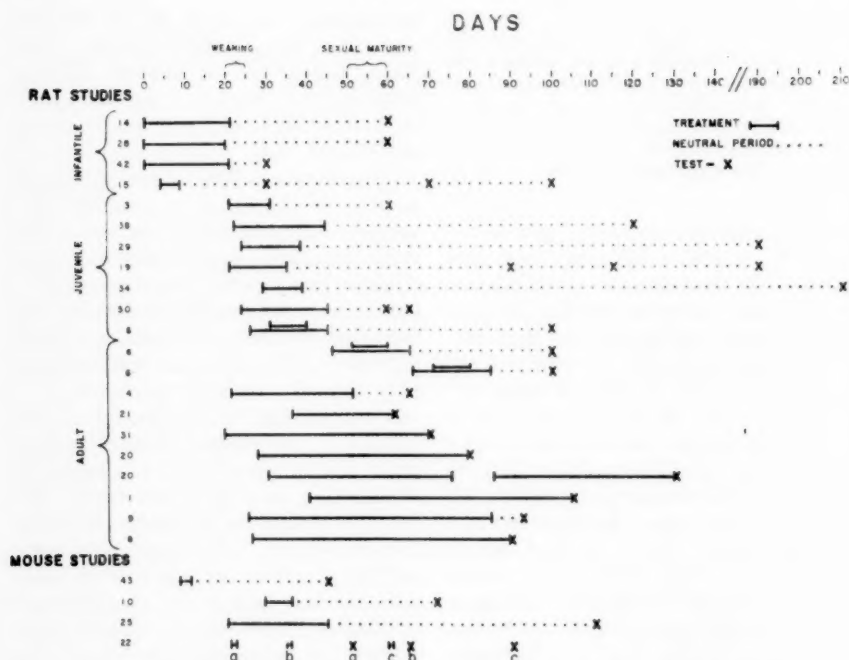
FIG. 1. SUBJECTS' AGES AT TREATMENT AND TEST IN SELECTED EXPERIMENTAL STUDIES

41) were taken during the treatment period and those experiments in which the tests began before or shortly after the rats reached sexual maturity (14, 15, 21, 26, 28, 42, 50). Those experiments, in which the animals are tested at intervals during the treatment, aid in the interpretation of adult behavior but are essentially descriptions of the maturational process.

The interval between the end of the treatment period and the beginning of the test period also varies considerably throughout the experiments. Some test periods began immediately after the treatment (8, 20, 21, 31) and, therefore, test for the *immediate* effects of the experience rather than *later* effects. Many treatments, of course, necessitate the immediate testing for effects in order to show the responses of an animal when first permitted to see or when first released from restriction (5, 36, 44, 45). Another group of experiments includes an interval of varying length between the early treatment and the later tests (4, 6, 9, 10, 14, 26, 28, 29). This interval has been as short as 5 days (9) and as long as 150 days (29). During this interval, both experimental and control groups are usually treated alike. This period of "neutral" experience for experimental and control animals is essential for testing the effect of early experiences upon adult behavior.

The duration of the testing period is necessarily varied according to the

type of performance task required of the animal. Some testing procedures have lasted almost as long as two years (42) while others have terminated after one week (21). Although the longer periods may contribute to our knowledge of persistence, which will be discussed later, they may also increase the danger of confusing the experiences gained early in life with those gained during the testing procedures. In general, the length of the testing period should be as short as possible in order to reduce practice effects. The minimum length which will give reliable results is probably the optimum.

DURATION OR QUANTITY OF EARLY EXPERIENCE

The hypotheses concerning the duration or quantity of the early experience propose that the experience may be anywhere from a brief period of exposure, such as a traumatic experience (12) or imprinting (30), to a relatively long accumulation of minor perceptual cues, such as learning to perceive (16). Most investigators have given their subjects relatively long periods of treatments, perhaps less because of a theoretical orientation than because they were eager to include a significant developmental period within the treatment span. The shortest treatment periods have been 1 day (10, 22) and 4 days (15, 43). Long treatment periods have lasted for as many as 65 days (1, 8). The short treatment periods provide a valid method for testing the effect of trauma and for discovering critical periods. While long treatment periods will do little to establish critical periods, they are useful for studying the effects of a particular kind of treatment on a particular performance task.

The most successful method for locating critical periods in the life span of an animal is to vary the age at treatment and vary the duration of the treatment. Hymovitch (20) has adequately demonstrated the technique of this procedure, and Eingold (6) has systematized it in a recent doctoral thesis. In the latter's study six experimental groups of rats were exposed to a free-environment treatment in a factorial design with two different lengths of exposure in each of three age classes. This procedure permitted Eingold to draw conclusions concerning the effects of the treatment with different durations of exposure as well as the effects at different ages. With a different procedure, King (24) found a 25-day treatment given to mice immediately after they were weaned had an effect on adult behavior. By reducing the duration of the treatment to 10 days and then to 5 days in subsequent groups, he was able to discover the minimum amount (duration) of treatment necessary to produce the effect on the adult mice. This amount of treatment was then given to another group of mice at adulthood. Fredericson (10) used a similar technique, by reducing the treatment from 7 days to 1 day; however, he did not give the treatment to mice at different ages.

If the age-location and duration of a critical period can be established, physiological and anatomical changes during this period can be correlated with the effects of the experiences. Physiological correlations with behavior may not enable the isolation of the effective variable because physiological and anatomical stages in maturation "cannot usually be controlled independently of each other" (2, p. 257). However, correlations comparing several strains or species

may eventually reduce the importance now given to age per se by permitting, for example, experimental manipulation at the period of myelinization, or at a given period of sensory sensitivity, or after a certain proportion of cholinesterase in the brain has been attained.

TYPE OR QUALITY OF EXPERIENCE

The types of experiences given to the developing subjects cover such a wide array that any enumeration without some classification is meaningless. On the other hand, no single classification adequately eliminates overlapping. Most experiments on this subject were stimulated by the theories of either Freud or Hebb (2). "The impact of Freudian theory, which led to various studies, such as those which limited food supply and feeding responses in young animals" (2, p. 239) are usually attempts to arrest psychosexual development of the subjects (11) or to measure later development of food directed motivation (18, 19, 29, 33, 34, 42). Experiments involving stress, whether shock (14, 28, 38, 43), temperature extremes, rotation, or intense auditory stimulation (14, 15) are usually designed to provide a traumatic experience to the developing organism (12). A second group of treatments which reduce stress through handling (3, 26, 28, 34, 38) or gentling (50, 51) may be placed in the same category. Experiments derived essentially from Hebbian concepts usually modify the environment of the subject. This modification may be either to increase the environmental complexity (4, 6, 9, 31) or to restrict the environment of the subject (4, 7, 20, 44, 45) beyond that of the usual conditions under which the animal is raised. The third group of treatments may fit either of

the two theories, but probably deserves a separate category because the social environment is manipulated. This manipulation may be in the form of isolation (25, 48) or exposure of the young subject to an aggressive (22) or sexual (21) situation.

Within each category the variety of treatments that has been used is extensive (Fig. 2). Electric shock, for example, may include different intensities, durations, and frequencies in each experiment. A restricted environment may be stovepipes (20), squeeze boxes (4), or small cages (7, 44). Visual experience has been manipulated by blinding (20), raising animals in darkness (36), covering eyes with opaque or prismatic lenses (17), or enclosing the subject in a transparent cage (9). Social treatments include attacks by an adult (22), fighting over food (10), exposure to the opposite sex (21), or living together in a group (25).

This variety of early treatments has been necessary in order to investigate the scope of treatments which have an effect upon adult behavior and to test the hypotheses of the different theories. Further contributions to the effect of early experiences on adult behavior will probably depend less on devising new types of treatments than upon examining the more "important theoretical issue . . . of the relationship between the quality of the early experience and the nature of the task to be solved" (9, p. 220).

TYPES OF PERFORMANCE TASK

After the animal has been given any of the preceding experiences, it is subjected to some test in order to learn the effect of the experience. Although many different tests are used, they are not as heterogeneous as the

PERFORMANCE TASK

EARLY EXPERIENCE		LEARNING				EMOTION			CONSUM.		SOCIAL	
		MAZE	DISCRIM.	CONDIT.	OTHER	ELIMINATE	ACTIVITY	AUD. STIM.	HOARDING	FOOD CONSUMP.	AGGRESSION	SEX
STRESS	SHOCK	14	14	28 43		14 43	38	14				
	AUD. STIM.	14	14			14 15	15	14				
	OTHER	14	14			14		14				
	HANDLING		3	28 26		26 38 51	50 51			38 51		
	FOOD DEPRIV.			29			29		18 19 33	19 34		
ENVIRON.	FREE MOTOR	4,6 8,9 20	4		31	31	31	4				
	FREE VISUAL	9 20										
	RESTRICTED	4,8 20	4									
SOCIAL	AGGRESSION										10 22	
	SEX											21
	GROUP					42	42	42	42	42	25 42	1

FIG. 2. TYPES OF EARLY EXPERIENCE AND PERFORMANCE TASK OF SUBJECTS IN EXPERIMENTS

treatments (Fig. 2). For the most part, the tests fall into four categories: emotion, learning, consummatory responses, and social. Emotional tests are most frequently activity and defecation scores in an open field (**15, 51**). Occasionally the field is electrified (**38**) or an intense auditory stimulus is applied (**4, 14, 42**) to increase emotional responses. Learning tests are usually either mazes (**4, 9, 14, 20**) or avoidance conditioning (**26, 28, 43**). The Lashley jumping tests (**4, 14**), a discrimination box (**3**), and operant conditioning (**29**) have also been used. Consummatory responses are chiefly those of hoarding (**18, 19, 33, 42**), food consumption (**34, 42**), and water consumption (**27**). The social tests are principally those developed for sex (**21, 49**) or aggression (**10, 22, 25**).

The number of tests given to any single group of subjects varies from 1 to 10 (**42**). Although most investigators administered only 1 test, the mean for 16 of the experiments reviewed here is 2.4 tests. This does not include training or extinction measures as separate tests, although both have been reported as separate indices in the results. In experiments including more than one test, the tests usually follow each other instead of being given in a random order or such a systematic order as provided by a latin square. This multiple testing procedure raises the question of what effect prior tests have upon subsequent ones. If one test affects the performance of the subject in another test, the direct effect of the early experience becomes lost in the complexity of interactions. One treatment and one kind of test, with a number of measures, is the most satisfactory method unless the investigator is specifically searching for interactions. Some animals, such

as dogs and monkeys, are so expensive to rear that it is too costly to offer the grown subject only one test. With these animals, which may be particularly affected by preceding tests, every procedure should include a schematic or random order of testing and/or an examination of interactions among the tests and treatment.

The results of experiments incorporating more than one test may yield important information regarding the extent to which a certain type of early experience may generalize into different patterns of behavior. Does a complex environment early in life affect adult emotionality as well as adult learning? Although it is preferable to answer this question with two separate groups which have been given the same treatment, one group of mice showed that early experience apparently affected both their adult aggressive and sexual behavior, but these adult behavior patterns were not correlated (**23**). Such results may indicate which types of adult behavior patterns are affected by a single type of early experience. Additional factors related to the number of tests a subject receives are considered in the next section.

PERSISTENCE OF EFFECTS

One concept explicitly stated in the hypothesis concerning the effects of early experience is that the effects last throughout life (**2**) or are even irreversible (**30**). Despite the significance of these hypotheses, there have been few experimental tests of the persistence of the effects. Three different methods have been used by the investigators who have attempted to measure persistence.

The first method can be illustrated by Hall's (**15**) experiment. After subjecting 4- to 7-day-old mice to an intense sound, he counted the fre-

quency of their elimination at 30 to 40 days in an open-field washtub and then he repeated the same test when the mice were 100 to 110 days old. The first and second tests were *interrupted* by a period of 60 days, during which the mice were kept in a neutral environment. This procedure runs the risk of having the first test affect the performance of the animal during the second test. The second procedure has been to make a *continuous* series of observations after the initiation of a single test (27). For example, Levine (27) measured water consumption in a control and experimental group 4 consecutive days after 18 hours of water deprivation each day. The initial differences in the two groups disappeared by the fourth day. Other investigators have left the animals in a test situation for extended periods and then they have made additional observations from time to time (7, 39). Fisher (7), for example, put young dogs back with their litter mates after they were raised in isolation. He found that the dogs raised in isolation were first highly responsive to their group-raised litter mates, but later tended to withdraw and remain isolated from the litter mates over a period of several months. It is possible that the withdrawal of the isolated dogs from the litter mates was enhanced in this situation by compounding the effects of the early experience of isolation with the later experiences derived while living together with the other dogs. The third procedure, which may be considered a test of persistence, consists in giving two groups of animals the same treatment at different ages (6, 20). If the group which received an early treatment is different from the group receiving a later treatment, the effect of the early treatment can be said to persist

through the later period, because the later period of treatment does not produce the same effect. Since only one test is given to each group of animals in this procedure, a real measure of persistence is lacking. A single test only defines the type of behavior elicited. A second test given to the same group of animals at a later age will reveal whether the behavior defined in the first test persists until the second test. A fourth possible procedure, which has not been used, would be to give the same treatment to several groups of animals at the same age and test them at different adult ages. Such an experiment should be designed to isolate and evaluate all interactive effects.

Certainly further studies on persistence of the effects of an early experience are needed. Most procedures run the risk of confounding the variables of early and late experiences, because the later experiences may enhance or inhibit the effect of the early experience. However, mammalian subjects cannot be kept in a vacuum, and it is likely that canalization may occur (35). Procedures designed to interpret this type of interaction between the later experiences and the early experiences of the subject may ultimately contribute more to the understanding of how adult behavior patterns are formed than procedures which prevent any interaction from occurring by keeping subjects in a "neutral" environment for different periods of time before testing them. Until such procedures are developed, the most reliable experiments will include a neutral period.

GENETICS AND EARLY EXPERIENCE

The seventh and last variable considered here is the effect which different genetic backgrounds have upon subjects undergoing the same early

experience. Genetic effects may be illustrated either by comparison of different species of animals or by the study of breeds or strains within one species. Species comparisons make possible an evaluation of the generalizations made from one species to another. Breed and strain comparisons indicate the effectiveness of a particular experience operating upon different genotypes.

The effect of early experience on later behavior has been studied in only a few common laboratory mammals. No single study has made direct species comparisons using similar techniques for more than one species. As indicated earlier, most work has been done on the laboratory rat. A few studies have been done on inbred mice (**10, 15, 22, 25, 43**). Valenstein and co-workers (**13, 46, 47, 48, 49**) have undertaken an investigation of guinea pigs, including the effects of early experience upon adult sexual behavior. Published investigations on carnivores are limited to dogs (**5, 7, 44, 45**), although some work has been done on cats at the American Museum of Natural History. Experimental studies on primates include Riesen's (**36**) investigation of rearing chimpanzees in the dark and work currently being undertaken at the University of Wisconsin on rhesus monkeys.

Species comparisons provide a basis for generalizations from one species to another, but they fail to indicate the relation of early experience to different genotypes because of the wide variation in sensory-motor capacities and rates of maturation among different species. The effect of early experience on animals of different genetic backgrounds is best studied in closely related animals, such as breeds or strains. In such studies, animals of different genotypes are given the same treatment and the same tests which offer comparative data for assessing the importance of genetic factors. Similar early experiences have been shown to have different effects on the adult behavior of two inbred strains of guinea pigs (**47, 48**), two inbred strains of mice (**24**), and several breeds of dogs (**7**). The demonstration of strain or breed differences is only the first step in a genetic analysis, but it is an important first step because it answers the question: do organisms of different genetic backgrounds respond alike to the same early experience?

SUMMARY AND CONCLUSIONS

This review has attempted to examine some of the variables involved in the experiments dealing with the effects of early experience on later behavior. In order to accomplish this analysis, the methods used in the experiments on this subject were examined in relation to seven variables. The experiments as a whole reveal a lack of attention to variables other than the one being manipulated. Other variables were controlled as they suited a particular experimental design rather than as a possible important contribution to the effects of early experience. This is particularly obvious upon inspection of Fig. 1, where the age variables are illustrated. The age of the animal at the time of treatment, for example, has included practically any age prior to the time the later test was given. Consequently, we are unable to make any definitive statement concerning which ages are most affected by a given experience. One of the most effective procedures for studying this age variable is that employed by

Eingold (6), who gave the same experience to several groups of rats at different ages during their development.

The age at the time of testing also requires further exploration, both in respect to the absolute age at which the animals are tested and in relation to the age of the early experience. No experiment has attempted to test animals at different ages, with the possible exception of experiments repeating a test on the same group of animals. Until this variable is adequately examined experimentally, it is possible to conclude only that the early experience affects an animal at *some* later age.

In the examination of the duration or quantity of experience given to the animal, we again encounter too much diversity to permit us to conclude how long the animal must endure its early treatment. The model for future experiments in testing this variable is again best demonstrated by Eingold, who varied the duration of treatment from 10 to 20 days in three groups of rats at different ages (Fig. 1). This method has the advantage of keeping the age variable separate from duration. More refinement in this procedure could be made by including other groups of different durations at each age and by controlling the length of the neutral period. The intensity of the experience also deserves attention and may be easily controlled with some types of stimuli, such as shock. Since the quantity of the early experience is important from the clinical point of view, it deserves much more attention than it has received in most experiments.

The type or quality of experience has been explored more thoroughly than the other variables. Most experiments have included at least two types of treatment (control and experimental) in the experimental design while the other variables are controlled. Because most investigators have used type of experience as their independent variable and because a number of them have controlled the age of treatment in nearly the same fashion, some generalizations are warranted. At the present time there is no adequate theoretical review of this material beyond the few introductory remarks preceding the report of each experiment.

The types of performance tasks, or the dependent variable, used to test the later behavior, are sufficiently varied to enable an assessment of the best measures of the effect of the treatment, except that the same animals are usually given a number of different tests at different ages. This procedure confounds the type of performance task with the age at time of testing and exposes the subjects to the effect of one test upon the other tests. The obvious solution here is to use different groups of animals for each test and to hold the age at testing constant. Although this procedure may require many more subjects, the refinement of the performance task is important to any conclusions regarding how one type of early experience may generalize to different types of behavior.

Persistence of the effects is probably the most difficult variable to examine because of the possibility that prior tests may affect later tests. This review described three procedures which have been used and suggested a fourth untried procedure. The particular virtue of the fourth procedure is that it would permit evaluation of the interactive effects resulting from successive testing. Perhaps in this indirect manner we can pro-

vide sufficient evidence to permit some conclusions regarding the persistence of behavior.

The last variable—that of genetics—has been held constant, at least to the species level, in all experiments. However, only the experiment of Valenstein et al. (**48**) has attempted to manipulate this variable in a single experiment. Further experiments using slightly different genetic strains (preferably inbred strains) will probably reveal more about the effect of this variable than species comparisons. Ultimately it will be desirable to make species comparisons, particularly if equivalent treat-

ments and tests can be developed. At the present time any extrapolations beyond the species studied will be hazardous.

Until the effects of the seven variables discussed in this review are analyzed further, it is possible to accept only the general hypothesis that some early experiences affect later behavior. Once each variable is isolated and we can accurately predict its effect upon later behavior, we will be in a position to study interactions and ultimately arrive at a comprehensive theory to account for many patterns of adult behavior on the basis of the experiences of the developing subject.

REFERENCES

1. BEACH, F. A. Comparison of copulatory behavior of male rats raised in isolation, cohabitation, and segregation. *J. genet. Psychol.*, 1942, **60**, 121–136.
2. BEACH, F. A., & JAYNES, J. Effects of early experience upon the behavior of animals. *Psychol. Bull.*, 1954, **51**, 239–263.
3. BERNSTEIN, L. A note on Christie's: "experimental naïveté and experiential naïveté." *Psychol. Bull.*, 1952, **49**, 38–40.
4. BINGHAM, W. E., & GRIFFITHS, W. J. The effect of different environments during infancy on adult behavior in the rat. *J. comp. physiol. Psychol.*, 1952, **45**, 307–312.
5. CLARKE, R. S., HERON, W., FETHERSTON-HAUGH, M. L., FORGAYS, D. G., & HEBB, D. O. Individual differences in dogs: preliminary report on the effects of early experience. *Canad. J. Psychol.*, 1951, **5**, 150–156.
6. EINGOLD, B. Problem-solving by mature rats as conditioned by the length, and age at imposition, of earlier free-environment experience. Unpublished doctor's dissertation, Univer. of Florida, 1956.
7. FISHER, A. E. The effects of differential early treatment on the social and exploratory behavior of puppies. Unpublished doctor's dissertation, Pennsylvania State Univer., 1955.
8. FORGAYS, D. G., & FORGAYS, J. W. The

nature of the effect of free-environmental experience in the rat. *J. comp. physiol. Psychol.*, 1952, **45**, 322–328.
9. FORGUS, R. H. Early visual and motor experience as determiners of complex maze-learning ability under rich and reduced stimulation. *J. comp. physiol. Psychol.*, 1955, **48**, 215–220.
10. FREDERICSON, E. Competition: The effects of infantile experience upon adult behavior. *J. abnorm. soc. Psychol.*, 1951, **46**, 406–409.
11. FREUD, S. Three essays on the theory of sexuality. In *The basic writings of . . . .* New York: Modern Library, 1938.
12. FREUD, S. The aetiology of hysteria. In *Collected papers.* London: Hogarth Press, 1950. Vol. 1.
13. GOY, R. W., & YOUNG, W. C. The importance of genetic and experimental factors for organization of sexual behavior patterns in the female guinea pig. *Anat. Rec.*, 1956, **124**, 296. (Abstract)
14. GRIFFITHS, W. J., & STRINGER, W. F. The effects of intense stimulation experienced during infancy on adult behavior in the rat. *J. comp. physiol. Psychol.*, 1952, **45**, 301–306.
15. HALL, C. S., & WHITEMAN, P. H. The effects of infantile stimulation upon later emotional stability in the mouse. *J. comp. physiol. Psychol.*, 1951, **44**, 61–66.
16. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.

17. HESS, E. H. Maturation and learning in the development of pecking accuracy in chicks. *Amer. Psychologist*, 1953, **8**, 367. (Abstract)

18. HUNT, J. McV. The effects of infant feeding-frustration upon adult hoarding in the albino rat. *J. abnorm. soc. Psychol.*, 1941, **36**, 338–360.

19. HUNT, J. McV., SCHLOSBERG, H., SOLOMON, R. L., & STELLAR, E. Studies of the effects of infantile experience on adult behavior in rats. I. Effects of infantile feeding frustration on adult hoarding. *J. comp. physiol. Psychol.*, 1947, **40**, 291–304.

20. HYMOVITCH, B. The effects of experimental variations on problem solving in rats. *J. comp. physiol. Psychol.*, 1952, **45**, 313–320.

21. KAGAN, J., & BEACH, F. A. Effects of early experience on mating behavior in male rats. *J. comp. physiol. Psychol.*, 1953, **46**, 204–208.

22. KAHN, M. W. The effect of severe defeat at various age levels on the aggressive behavior of mice. *J. genet. Psychol.*, 1951, **79**, 117–130.

23. KING, J. A. Sexual behavior of C57BL/10 mice and its relation to early social experience. *J. genet. Psychol.*, 1956, **88**, 223–229.

24. KING, J. A. Relationships between early social experience and adult aggressive behavior in inbred mice. *J. genet. Psychol.*, 1957, **90**, 151–166.

25. KING, J. A., & GURNEY, N. L. Effect of early social experience on adult aggressive behavior in C57BL/10 mice. *J. comp. physiol. Psychol.*, 1954, **47**, 326–330.

26. LEVINE, S. A further study of infantile handling and adult avoidance learning. *J. Pers.*, 1956, **25**, 70–80.

27. LEVINE, S. The effect of shock and handling in infancy on consummatory response in adulthood. *Amer. Psychologist*, 1956, **11**, 397. (Abstract)

28. LEVINE, S., CHEVALIER, J. A., & KORCHIN, S. J. The effects of early shock and handling on later avoidance learning. *J. Pers.*, 1956, **24**, 475–493.

29. LITTMAN, R. A. Infantile experience and adult behavior in the white rat. *J. genet. Psychol.*, 1956, **88**, 11–24.

30. LORENZ, K. Der Kumpan in der Umwelt des Vogels. *J. Orn., Lpz.*, 1935, **83**, 137–213; 289–413.

31. LUCHINS, A. S., & FORGUS, R. H. The effect of differential post-weaning environment on the rigidity of an animal's behavior. *J. genet. Psychol.*, 1955, **86**, 51–58.

32. MCCLELLAND, W. J. Differential handling and weight gain in the albino rat. *Canad. J. Psychol.*, 1956, **10**, 19–22.

33. MCKELVEY, R. K., & MARX, M. H. Effects of infantile food and water deprivation on adult hoarding in the rat. *J. comp. physiol. Psychol.*, 1951, **44**, 423–430.

34. MARX, M. H. Infantile deprivation and adult behavior in the rat: retention of increased rate of eating. *J. comp. physiol. Psychol.*, 1952, **45**, 43–49.

35. MURPHY, G. *Personality.* New York: Harper, 1947.

36. RIESEN, A. H. The development of visual perception in man and chimpanzee. *Science*, 1949, **106**, 107–108.

37. ROSS, S. Sucking frustration in neonate puppies. *J. abnorm. soc. Psychol.*, 1951, **46**, 142–149.

38. SCOTT, J. H. Some effects at maturity of gentling, ignoring or shocking rats during infancy. *J. abnorm. soc. Psychol.*, 1955, **51**, 412–414.

39. SCOTT, J. P. Social behavior, organization and leadership in a small flock of domestic sheep. *Comp. Psychol. Monogr.*, 1945, **18**, No. 4.

40. SCOTT, J. P., FREDERICSON, E., & FULLER, J. L. Experimental exploration of the critical period hypothesis. *Personality*, 1951, **1**, 162–183.

41. SCOTT, J. P., & MARSTON, M. V. Critical periods affecting the development of normal and mal-adjustive social behavior of puppies. *J. genet. Psychol.*, 1950, **77**, 25–60.

42. SEITZ, P. The effects of infantile experience upon adult behavior in animal subjects: I. Effects of litter size during infancy upon adult behavior in the rat. *Amer. J. Psychiat.*, 1954, **110**, 916–927.

43. STANLEY, W. C., & MONKMAN, J. A. A test for specific and general behavioral effects of infantile stimulation with shock in the mouse. *J. abnorm. soc. Psychol.*, 1956, **53**, 19–22.

44. THOMPSON, W. R., & HERON, W. The effects of early restriction on activity in dogs. *J. comp. physiol. Psychol.*, 1954, **47**, 77–82.

45. THOMPSON, W. R., & HERON, W. The effects of restricting early experience on the problem-solving capacity of dogs. *Canad. J. Psychol.*, 1954, **8**, 17–31.

46. VALENSTEIN, E. S. The role of learning and androgen in the organization and

display of sexual behavior in the male guinea pig. *Amer. Psychologist*, 1954, **9**, 486. (Abstract)

47. VALENSTEIN, E. S., RISS, W., & YOUNG, W. C. Sex drive in genetically heterogeneous and highly inbred strains of male guinea pigs. *J. comp. physiol. Psychol.*, 1954, **47**, 162–165.

48. VALENSTEIN, E. S., RISS, W., & YOUNG, W. C. Experiential and genetic factors in the organization of sexual behavior in male guinea pigs. *J. comp. physiol. Psychol.*, 1955, **48**, 397–403.

49. VALENSTEIN, E. S., & YOUNG, W. C. An experiential factor influencing the effectiveness of testosterone propionate in eliciting sexual behavior in male guinea pigs. *Endocrinology*, 1955, **56**, 173–177.

50. WEININGER, O. Physiological damage under emotional stress as a function of early experience. *Science*, 1954, **119**, 285–286.

51. WEININGER, O. The effects of early experience on behavior and growth characteristics. *J. comp. physiol. Psychol.*, 1956, **49**, 1–9.

EXACT PROBABILITY TREATMENTS OF FACTORIAL DESIGNS

JEROME L. MYERS

University of Massachusetts

The research psychologist is often confronted by data which represent the frequency of use of various response categories. This occurs when rating scales are employed, or when the Ss' responses are categorized in one of a number of ways. Sutcliffe (5) has recently presented a χ^2 method for analyzing the data when a factorial design is used. This paper will describe alternative approaches of an exact probability nature.

A TEST FOR RANDOM SAMPLES

Freeman and Halton (4) have derived an exact probability approach, applicable whenever (*a*) either the parent population is of infinite size or the sampling is done with replacement of the sampled members, and (*b*) the sampling is random. Table 1 presents a $2 \times 2 \times 2$ design. The number of individuals falling into each cell is indicated by m_{ijk} where i indicates the row, j the column, and k the block. The probability of the observed distribution of frequencies is given by

$$\sum_k \sum_j m_{1jk}, \quad \sum_k \sum_j m_{2jk}, \quad \sum_k \sum_i m_{i1k},$$

$$\sum_k \sum_i m_{i2k}, \quad \sum_j \sum_i m_{ij1}, \quad \sum_j \sum_i m_{ij2}.$$

TABLE 1
ILLUSTRATION OF A $2 \times 2 \times 2$ DESIGN

	A_1		A_2	
	B_1	B_2	B_1	B_2
C_1	m_{111}	m_{121}	m_{112}	m_{122}
C_2	m_{211}	m_{221}	m_{212}	m_{222}

Use Formula 1 to assess the probability of each of these tables. The probabilities so obtained must of course sum to 1.00. The significance level is obtained by summing the probability of the observed table with all those probabilities which are equal to or less than it.

This approach can be extended to any number of independent classifications and levels. In general, (*see formula at top of page 60*).

$$\frac{\left(\sum_k \sum_j m_{1jk}\right)!\left(\sum_k \sum_j m_{2jk}\right)!\left(\sum_i \sum_k m_{i1k}\right)!\left(\sum_i \sum_k m_{i2k}\right)!\left(\sum_i \sum_j m_{ij1}\right)!\left(\sum_i \sum_j m_{ij2}\right)!}{\left\{\left(\sum_i \sum_j \sum_k m_{ijk}\right)!\right\}^2 m_{111}!m_{121}!m_{211}!m_{221}!m_{112}!m_{122}!m_{212}!m_{222}!} . \quad [1]$$

The manner of obtaining the significance level will now be described. Write out all possible $2 \times 2 \times 2$ tables having the observed marginal totals as follows:

where N equals the total of all the cell frequencies and r equals the number of classifications. The significance level is again computed in the manner described above.

59

$$p = \frac{\text{(the product of the factorials of the marginal totals)}}{(N!)^{r-1} \text{(the product of the factorials of the cell frequencies)}}$$

A Test for the Mixed Case

In the usual experimental situation, a fixed number of subjects are assigned to each combination of treatment levels, and only the response categories are randomly sampled. Consider Table 1 again, assuming that there are a fixed number of subjects in cells A_1B_1, A_1B_2, A_2B_1 and A_2B_2 and that C_1 and C_2 correspond to two response categories. To ascertain the significance level of the AC interaction retabulate the data in the form of an $A \times C$ table, ignoring the B classification. Fisher's exact probability test (2) may now be applied. Fisher's test and that of Freeman and Halton are the same in the case of the 2×2 contingency table. The BC interaction may be assessed in a similar manner. If there are more than two levels of either factor entering into the two-way table, the Freeman and Halton test should be used. Higher order interactions (e.g. ABC) must be somewhat differently dealt with, in the case of the mixed design. The remainder of this paper will describe a procedure for dealing with such problems.

There appear to be three restrictions which dictate our choice of a test for higher order interactions: (*a*) the marginal totals are considered fixed; (*b*) the total frequency for each combination of classification levels (exclusive of the response classification) is considered fixed; and (*c*) the first order interactions are considered as fixed effects. The second and third restrictions differentiate the test from Freeman and Halton's test for the random sample. The third restriction dictates a procedure which would be analogous to subtracting χ^2_{AC} and

χ^2_{BC} from χ^2_{total} in Sutcliffe's test.

Since the test is more easily executed than described, an illustration will be presented. Consider the data in Table 2. Since the AB cell totals are fixed (they have been made equal for simplicity, but this is not neces-

TABLE 2

THE OBSERVED FREQUENCY DISTRIBUTION IN THE EXAMPLE

	A_1			A_2		
	B_1	B_2		B_1	B_2	
C_1	4	0		3	5	12
C_2	4	8		5	3	20
	8	8		8	8	32

sary) the number of ways in which the observed distribution of C_1 and C_2 could have occurred is given by $\binom{8}{4}\binom{8}{0}\binom{8}{3}\binom{8}{5} = 219{,}520$. The question is, In how many ways could 32 responses be distributed, given the restrictions that: there are 8 responses in each of the four AB cells (restriction (*b*) above); there are 16 in A_1, in A_2, in B_1, and in B_2, and 12 in C_1 and 20 in C_2 (restriction *a* above); and the frequency in A_1C_1 is 4, in A_2C_1 is 8, in B_1C_1 is 7, in B_2C_1 is 5, etc. (restriction *c* above). We could have obtained the distribution in Table 3

TABLE 3

AN ALTERNATIVE FREQUENCY DISTRIBUTION UNDER THE RESTRICTIONS

	A_1			A_2		
	B_1	B_2		B_1	B_2	
C_1	0	4		7	1	12
C_2	8	4		1	7	20
	8	8		8	8	32

under these restrictions. The number of ways in which that could have occurred is $\binom{8}{0}\binom{8}{4}\binom{8}{7}\binom{8}{1} = 4,480$. If we enumerate all possible tables we have

$$\binom{8}{0}\binom{8}{4}\binom{8}{7}\binom{8}{1} = 4,480$$

$$\binom{8}{1}\binom{8}{3}\binom{8}{6}\binom{8}{2} = 351,232$$

$$\binom{8}{2}\binom{8}{2}\binom{8}{5}\binom{8}{3} = 2,458,624$$

$$\binom{8}{3}\binom{8}{1}\binom{8}{4}\binom{8}{4} = 2,195,200$$

$$\binom{8}{4}\binom{8}{0}\binom{8}{3}\binom{8}{5} = \frac{219,520}{5,229,056}$$

Note that continued enumeration leads us to $\binom{8}{5}\binom{8}{-1}\binom{8}{2}\binom{8}{6}$. Negative frequencies are, however, meaningless, and in fact $(r_{-x}) = 0$ by mathematical convention (1, p. 40) so that the expression vanishes.

The significance level of the ABC interaction is

$$p = \frac{219,520 + 4,480}{5,229,056} = .042$$

In short, p equals the number of ways in which the observed event or one less likely could have occurred, divided by the total number of events possible, given the three restrictions

previously described. Sutcliffe's test yields a χ^2 of 4.79, significant between the .02 and .05 levels.

ADDITIONAL COMMENTS

This test can be simply generalized to any interaction for any number of classifications and levels, if the experimenter understands the three restrictions and their implications for the analysis. Note that if j cells are involved, and if there are i response categories in each, the number of ways of arriving at the observed table is

$$\frac{\left(\sum_1^i m_{i1}\right)!}{m_{11}!m_{21}!\cdots m_{i1}!} \frac{\left(\sum_1^i m_{i2}\right)!}{m_{12}!m_{22}!\cdots m_{i2}!}$$
$$\cdots \frac{\left(\sum_1^i m_{ij}\right)!}{m_{1j}!m_{2j}!\cdots m_{ij}!}. \quad [2]$$

The experimenter may find a table of logarithms of factorials helpful in the calculations (3, pp. 108–109).

The technique described is simple to compute, requires no table of significance values, and yields an exact probability estimate. The procedure should prove especially advantageous with small samples because of the further reduction in computational labors, and especially because of the limitations of χ^2 in dealing with small cell frequencies.

REFERENCES

1. FELLER, W. *An introduction to probability theory and its applications. Vol. I.* New York: Wiley, 1950.
2. FISHER, R. A. *Statistical methods for research workers,* (5th ed.). Edinburgh: Oliver & Boyd, 1934.
3. FISHER, R. A., & YATES, F. *Statistical tables for biological, agricultural and medical research* (4th ed.). New York: Hafner, 1953.
4. FREEMAN, G. H., & HALTON, J. H. Note on an exact treatment of contingency, goodness of fit, and other problems of significance. *Biometrika*, 1951, **38**, 141–149.
5. SUTCLIFFE, J. P. A general method of analysis of frequency data for multiple classification designs. *Psychol. Bull.*, 1957, **54**, 134–137.

CORRECTING PERSONALITY SCALES FOR RESPONSE SETS OR SUPPRESSION EFFECTS

HAROLD WEBSTER

Mellon Foundation, Vassar College

Recently there has been increased interest in how personality test scores are affected by subjects' general "tendency to agree" or "tendency to disagree" with the items (**1, 2, 5, 7**). There has also been a persistent concern with attempts by subjects to convey through their test responses either overly favorable or overly unfavorable impressions (**2, 3, 4, 6, 9, 10**). For some kinds of tests these two kinds of effects may be nearly indistinguishable, or at least so it would appear from a study by Fricke (**5**) and from some recent work at Vassar College using personality inventory data (**10**). The need for continuing research in this area seems more urgent than ever before because of a current popular skepticism about the merits of personality testing (**11**).

The present paper is not concerned with the psychological aspects of response set or suppression or plain lying by subjects who take personality tests, although it is obvious that these are of greatest importance to psychologists. It is instead concerned with the secondary (but nearly equally important) problem of correcting personality scale scores for such effects. Statistical method in this area seems largely to have been ignored. With few exceptions (**3, 7, 8**) investigators appear to have believed that empirical trials, in many cases with small samples, would be superior to more logical methods.

Consider a personality scale, yielding raw scores T, comprised of the typical kinds of true-false items of which, say, 91% are scored "true."

Cronbach (**2**) points out that there is much evidence that a subject's score on such a scale can be raised merely by his tendency to respond "true" (or else depressed by his tendency to answer "false"), a characteristic which undoubtedly varies from subject to subject. In order to correct the scores T for this effect, we first must be able to measure the effect. This is known to be easy; in fact, if one scores a large number (say 200) of randomly selected true-false personality test items "true," the total score t will then be found to have rather good reliability, and in addition it will, on logical and operational grounds, be a measure of "tendency to answer 'true' on the test yielding scores t." Of course, if it is desired, obtained reliability may then be increased further by item analyses (**10**).

It next follows that the correlation between T, the original uncorrected scale scores, and t, the scores for tendency to answer "true," is of immediate interest. Even if the regression should happen to be nonlinear, which is unlikely, this correlation could still be useful as a measure of the linear component actually present. We decide that what is needed is a new score T' for each subject; a score, which, unlike T, is independent of his response set score t. The simplest definition of such a score is

$$T' = T + kt, \qquad [1]$$

where k is determined such that t and T' represent independent measures, preferably for large representative samples. Now T' is *very likely* to be

completely independent of t if it is uncorrelated with it, and the condition for such a zero correlation is that the covariance

$$\text{Cov } (T', t) = 0. \qquad [2]$$

From [1] and [2]

$$\text{Cov } (T', t) = \text{Cov } (T, t) + kV(t) = 0, \quad [3]$$

where $V(t)$ is the variance of t. Solving [3] for k,

$$k = -\frac{\text{Cov } (T, t)}{V(t)} = -\frac{S_t}{S_T} r_{Tt}, \qquad [4]$$

where the Ss are standard deviations.

The required weight k for use in [1] is therefore the linear regression of t on T, so that the new score T', which is "uncontaminated" by response set t, can be written for the ith subject,

$$T'_i = T_i - b_{tT} t_i, \qquad [5]$$

where b_{tT} is this regression coefficient. Allen L. Edwards points out, in a personal communication, that since bt is a regression estimate of T, which may be written $\hat{T}$, the correlation between $(T - \hat{T}) = T'$ and t is necessarily zero because the variance of residual errors cannot contribute to it. He further notes that the cross products for the correlation between

T' and t may, referring to [5], be written

$$\sum t(T - \hat{T}) = \sum tT - b \sum t^2, \quad [6]$$

which is seen to equal zero when

$$\frac{\sum tT}{\sum t^2}$$

is substituted for b.

There seems to be no better way of correcting for "suppression" or for response set. If T and t are not reliable, then [5] will not work, but in this case neither will other methods. Some idea of the effect of such adjustments on the reliability of final scores may be obtained from the data of Table 1.

EXAMPLES

The statistics of Table 1 are for seven true-false personality scales which were developed over a period of five years as part of a general investigation of personality development during the college years (10). The data are for a standardization sample of 906 women undergraduate students, 453 freshmen and 453 seniors. With the exception of MF, the initial reliabilities r_{TT}, as measured by KR 21, are fairly high; reliability of MF, a traditional type of mascu-

TABLE 1

KR 21 RELIABILITIES, CORRELATIONS AND REGRESSION COEFFICIENTS USED TO ESTIMATE RELIABILITIES OF SCORES AFTER ADJUSTMENT FOR RS ($N = 906$)

Scale T	No. of Items	r_{TT}	$r_{(RS)T}$	$b_{(RS)T}$	$r^*_{T'T'}$
Social Maturity (SM)	149	.88	.35	.45	.85
Developmental Status (DS)	72	.89	− .29	− .54	.86
Impulse Expression (IE)	124	.89	− .58	− .80	.79
Dominance and Confidence (DC)	92	.87	.53	.93	.78
Social Integration (SI)	99	.92	.86	1 .39	.61
Masculine Role (MF)	96	.54	.14	.42	.56
Repression and Suppression (RS)	157	.93			

* Reliabilities of scales, estimated by Formula 7, after adjustment for effects measured by RS.

linity-femininity scale, would be higher if male subjects had also been included in the sample. *RS*, which was developed after much item analysis, is comprised of 157 items scored "false" and therefore constitutes a measure of response set "*t*" according to the previous definition; reversing scoring so that all items were scored "true" would of course simply change the signs of the correlation and regression coefficients of Table 1. The approximate reliabilities of the scales after adjustment for the effects measured by *RS* are given in the final column of the table. These reliabilities were estimated by

$$\tilde{r}_{T'T'} = \frac{r_{TT} + b^2 r_{(RS)(RS)} - 2br_{(RS)T}}{b^2 + 1 - 2br_{(RS)T}}, \quad [7]$$

using data of Table 1. More exact values could be obtained by computing the reliability of each scale using the individual adjusted scores T', computed as in [5].

The importance of correlation between *RS* and the scale to be corrected is illustrated in Table 1, for the more extreme the correlation, the greater is the decrease in reliability after adjustment of the original raw scores. The importance of high reliability for *RS* itself is obvious from [7]; it is interesting that there is a slight *increase* in reliability for the *MF* adjusted scores. Reliability remains high for some of the adjusted scales in Table 1, so *theoretically* a second adjustment might be made; for example, some scales might next be corrected for "judged social desirability" of item responses, *if* this could be measured by a reliable and relatively independent scale. The possibility of using more than one adjustment on a given scale has not been investigated empirically.

REFERENCES

1. BASS, B. M. Authoritarianism or acquiescence? *J. abnorm. soc. Psychol.*, 1955, **51**, 616–623.
2. CRONBACH, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt.*, 1950, **10**, 3–31.
3. EDWARDS, A. L. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *J. appl. Psychol.*, 1953, **37**, 90–93.
4. FORDYCE, W. E. Social desirability in the MMPI. *J. consult. Psychol.*, 1956, **20**, 171–175.
5. FRICKE, B. C. Response set as a suppressor variable in the OAIS and MMPI. *J. consult. Psychol.*, 1956, **20**, 161–169.
6. GOUGH, H. G. On making a good impression. *J. educ. Res.*, 1952, **46**, 34–42.
7. HELMSTADTER, G. C. Procedures for obtaining separate "set" and "content" components of a test score. In the Educational Testing Service Annual Report, 1955–1956. (Abstract)
8. HUMM, D. C., & HUMM, KATHRYN A. Compensations for subjects' response-bias in a measure of temperament. *Amer. Psychologist*, 1947, **2**, 305. (Abstract)
9. MEEHL, P. E., & HATHAWAY, S. R. The K Factor as a suppressor variable in the MMPI. In C. S. Welsh & W. C. Dahlstrom (Eds.) *Basic readings on the MMPI in psychology and medicine.* Univer. Minnesota Press, 1956.
10. WEBSTER, H., FREEDMAN, M., & SANFORD, N. Research Manual for the VC Attitude Inventory and VC Figure Preference Test. Vassar College: Mary Conover Mellon Foundation, 1957.
11. WHYTE, W. H., JR. *The organization man.* New York: Simon and Schuster, 1956.